ABSTRACT

MAJIKES, JACOB MICHAEL. DNA Origami Folding Pathways: Implications for Design, Thermodynamics, and Kinetics. (Under the direction of Dr. Thomas LaBean.)

DNA nanotechnology implements the predictable self-assembly rules of DNA, allowing the adaptation of DNA from a biological tool for storage of genetic information to a biomimetic structural nanomaterial. DNA has been employed to organize organic and inorganic materials, as well as to create both static and dynamic nanostructures. Aided by the low cost of arbitrary sequence DNA oligomer synthesis and robust conjugation chemistries, DNA has developed as a promising nanofabrication tool. While under biological conditions the formation and thermodynamics of DNA are well known, nanotechnology applications typically lie well outside of those conditions.

This dissertation presents a new scaffold (miniM13) for DNA nanostructures and three new protocols to probe the folding and formation of DNA nanostructures. Development of these novel techniques improves the molecular assembly toolkit to enable new and exciting experimental systems.

The new miniM13 workbench scaffold presented in this work reduces the upfront costs of scaffolded DNA nanostructures by 50%, and reduces material costs by 70%. The use of this smaller scaffold allows faster iteration through design/prototype cycles, and through its lower upfront costs, facilitates experiments which explore diverse nanostructure designs.

The implementation of effective baseline correction procedures on fluorescence data for DNA nanostructure assembly/disassembly significantly improves the utility of the method; measurements of hysteresis between melt and anneal data are reported. Use of appropriate baseline correction and photobleaching corrections provides a valuable tool for DNA nanostructure characterization, especially when used with the high throughput of real-time polymerase chain reaction (RT-PCR) equipment.

The development of competitive multishape anneals (one scaffold simultaneously folded by multiple staple sets) allows examination of systemic folding preferences via which sections of a structure ‘win’ over their competition and are present in the final structure. Using the rapid prototyping of miniM13 structures, three shapes are brought into competition by mixing their staple sets together before annealing. Competition between the circle, 3D 6-helix bundle, and triangle staple sets over the miniM13 scaffolds supports the mutual role of GC content and conformational entropy in nanostructure folding. Presented results support nucleation and growth as a mechanism for the formation of DNA nanostructures.

Finally, a protocol to “chemically freeze”, or quench, DNA nanostructure formation at different times during the folding process is presented. Given the difficulty and limitations of previous such techniques, the developed chemical quench provides an elegant solution for directly probing nanostructure folding pathways. Results of chemical quenching the standard Tall Rectangle nanostructure further implicate the dominance of conformational entropy in structure formation.

Collectively, the presented techniques provide useful tools for understanding the design/folding relations of DNA nanostructures. The implementation of these techniques on model systems collectively provide general insight into DNA nanostructure folding while also elucidating the role of conformational entropy in cooperative formation.

DNA Origami Folding Pathways: Implications for Design, Thermodynamics, and Kinetics

by
Jacob Michael Majikes

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Materials Science and Engineering

Raleigh, North Carolina

2016

APPROVED BY:

Dr. Thomas LaBean Committee Chair	Dr. Lew Reynolds
Dr. Melissa Pasquinelli	Dr. Yaroslava Yingling

DEDICATION

To the family, friends and coworkers whose presence and advice have kept me grounded throughout graduate school.

BIOGRAPHY

Jacob Majikes is a researcher who has a poor understanding of what to include in a biography, he was born on November 17, 1987 to John and Kim Majikes. He attended NC State University for his B.S. in Materials Science and Engineering and pursued his Ph.D in the same department under Thom Labean. He was also the second graduating class from the Centennial Middle School on NC State's campus.

There was much rejoicing on his graduation due the noticeably decreased number of puns being said in Engineering Building 1 after his departure.

ACKNOWLEDGMENTS

I would like to acknowledge the help of many researchers whose advice and guidance have made this work possible.

First and foremost, my advisor Thom LaBean, and my committee members, Lew Reynolds, Melissa Pasquinelli, and Yaroslava Yingling.

I would like to thank my lab members Abhijit Rangekar, Briana Vogen, David Courson, Ronnie Pedersen, Alex Marchi, Nhat Ngo, and Nicole Estrich for their advice and time.

I would also like to thank Lucas Ferraz and Hannah Fennel for their work as undergraduate researchers.

I thank Tim Sit, and Steve Lommel, for their advice and for the use of their RT-PCR equipment.

I thank Jean-Phillippe Sobczak, Juan Elezgaray, and Andre Estevez-Torres for their quick responses and technical advice. I also thank Mingdong Dong of Aarhus University, Kurt Gothelf of Aarhus University, Stanley Brown of Copenhagen University, and Jennifer Black of Oak Ridge National Lab for allowing me to access their equipment and expertise.

Finally, I thank my friends and family for keeping me sane enough to pursue this degree.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 4: Competitive Annealing and Multishape Nanostructures

CHAPTER 1: Introduction and Motivation

Section 1.1: Biomimetics and DNA Nanotechnology

In 1998, the annual change in world wide computing power, driven by top-down nanomanufacturing, peaked[1]. This led to a robust debate of whether, and when, Moore's law will cease to apply. In any case, it is accurate to say that the return on investment for research and new techniques in silicon electronics is not as high as it was twenty, or even ten, years ago. At the same time, our ability to synthesize and manipulate complex macromolecules has developed, and received particular spotlight[2]. Top-down nanofabrication has become a cornerstone of our technological lives, while a new frontier of bottom-up nanofabrication has opened[3].

Biological systems act in 3D, elegantly marry structure and function, respond actively to their environments, and are capable of replication and evolution. The lofty goal of mimicking these abilities is met with challenges of equal proportion. Despite ever-improving characterization tools, the need to capture 3D structure and motion of macromolecules hinders our ability to gather structural data. Despite developing conceptual understanding of biologic systems, our ability to parse the nuance present in these systems is outflanked by natural diversity and complexity. Despite increasing synthetic techniques for precise polymers, we can only begin to mimic the capabilities of the biological systems that create the food we eat, the cotton we wear, and the raw materials of most products we consume.

To build complex materials on the nanoscale, bottom-up techniques have sought to mimic biology[4]. These attempts at biomimetic technology have shown great promise. Of these techniques, we address one of the most exciting, DNA nanotechnology[5].

Thanks to the massive demand for arbitrary-sequence DNA in biological and medical applications, as well as the efficiency of phosphoramidite synthesis[6], it is possible to obtain

low cost single-strand DNA, or ssDNA[7]. Thanks to the extent of biological research, we have a strong understanding of how arbitrary ssDNA will assemble into more complex systems[6].

On this foundation DNA has developed as a nanomaterial. At its nascence, and through the early 2010s, the price of synthetic DNA dropped precipitously each year, as shown in Figure 1.1 reproduced from Carlson. However, since approximately 2010, due to the limits of automation for phosphoramidite synthesis, the price of synthetic DNA oligomers has merely kept pace with inflation.

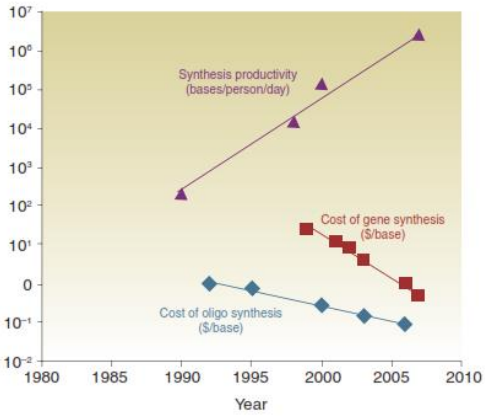


Figure 1.1: Synthetic DNA production costs to 2006[7]

Of the many considerations in the development of DNA nanotechnology, one which will recur through this work is that of design and cost. While there are many variations of DNA nanotechnology, we will focus on scaffolded nanostructures, or origami. We will discuss the details of scaffolded nanostructure assembly in section 1.4. For this introduction it suffices to describe scaffolded nanostructures as the combination of a long viral ssDNA "scaffold" and many short synthetic ssDNA oligomer "staples" which form an arbitrary predesigned structure[8]. These structures are limited only by imagination and the available scaffold length.

In implementing DNA nanostructures, two costs must be considered. The first is the material cost of the nanostructures, which will be vital for industry scale-up and implementation. The second is the upfront costs associated with design and minimum

purchases for creating a new structure, which limit innovation. Table 1.1 shows a brief summary of these costs, updated from the original Rothemund work[8]; details of cost estimates are discussed in chapter 2. Other types of DNA nanostructures, differ in details, the general trend is the same. The upfront costs are limited by the number of synthetic staples, and material costs by the scaffold.

Table 1.1: Example upfront and material costs of DNA nanostructures

	Upfront Costs		Material Costs	
	Min. Purchase	# Anneals	Cost / nMol	Cost / Anneal
Viral Scaffold	~$37	298	$9,322	$1.40
Synthetic Staples	~$1305	533,333	$198	$0.03
Design Cost	~$2,400	∞	-	-
Total	**$3,742**	**298**	**$9,520**	**$1.43**

For research purposes the per-anneal, or per standard sample for AFM imaging, costs are fairly reasonable. However, in order to iterate through multiple redesigns or to compare structure motifs, it is prohibitive. This is complicated by vastly different scales for research and for applications such as drug delivery or nanoelectronics. DNA nanostructures to deliver a single dose of chemotherapy could use as many as 41,000 anneals, or $60,000.

With the goals and costs of nanofabrication in mind, we will next discuss the myriad of applications toward which DNA nanotechnology has already progressed.

Section 1.2: Examples of DNA Nanotechnology

DNA as a nanomaterial is powerful in its molecular recognition. This can take the form of binding other molecules with aptamers[9], with strand displacement algorithms[10], or with the use of recognition as a tool for nanofabrication[8]. In this first function, DNA binds other nanomaterials, both organic and inorganic, through aptamers[9,11]. This is often performed by taking a library of aptamers, screening all of them to a nanomaterial or macromolecule of interest, using PCR to amplify aptamers that bound, then repeating the

process and sequencing the result. These binding techniques have much potential, and have been implemented towards numerous ends ranging from the coagulation cascade to chemotherapy.

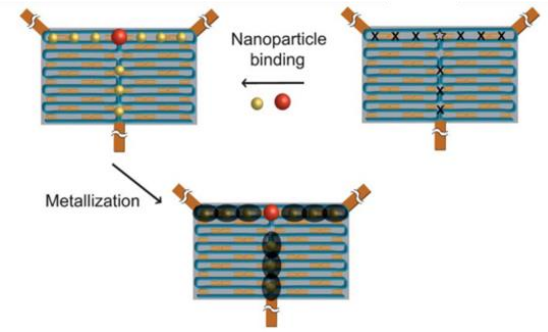


Figure 1.2: Schematic of single electron transistors patterned by DNA nanostructures

The use of DNA as a nanomaterial, which is of more relevance to this work, is as a programmable structural tool for nanofabrication. Structural DNA nanotechnology marries the low cost of synthetic single-strand DNA with the well-understood interactions between DNA and the diverse DNA conjugation toolkit[6,7,12]. This combination allows the creation of nearly arbitrary structures, decorated with a wide array of functional nanomaterials through conjugation chemistry. Figure 1.2 shows one such example, which uses thiol-DNA conjugation to label gold nanoparticles and nanorods to create the geometry of a single-electron transistor[13].

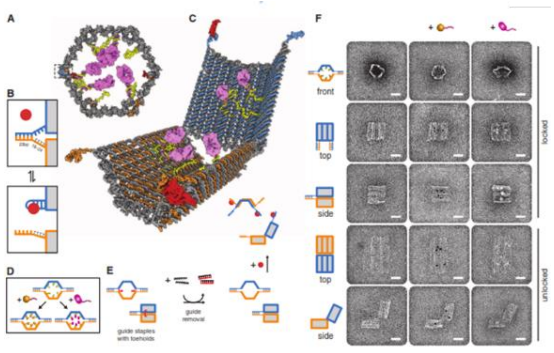


Figure 1.3: Nanostructure actuator to trigger cell apoptosis

This approach has yielded high-impact results. DNA has been used to contain and control the targeting of antibodies as per Church, reproduced in Figure 1.3[14]. DNA has been used to pattern multiplex assays, and to create plasmon resonance structures[15,16]. Synthetic pores which integrate into cellular membranes have been developed[17]. Algorithmic operations have been performed both with strand displacement[18,19] and with coupled fluorophores[20].

Numerous other advances have been made with structures other than origami. DNA tiles, grown in 3D, have created crystals which may be used to harvest light[21,22]. Small DNA structures can spatially control the conjugation of other nanomaterials[23], as shown in Figure 1.4, greatly improving our ability to address their locations.

Meanwhile, much effort has been expended to improve the processing capabilities of DNA as a nanomaterial. Through careful gradient centrifugation, the Liddle lab has significantly improved the yield during purification of decorated nanostructures[24]. The Dietz, Estevez-Torres, and Turberfield labs have all contributed significantly to the range of annealing protocols. The compatibility of DNA nanostructures with mixed solvents[25] and free drying[26] have been tested and confirmed. Additionally, alternate DNA assembly techniques such as the DNA brick[27] and the DNA single strand tile[28] have replaced the viral scaffold

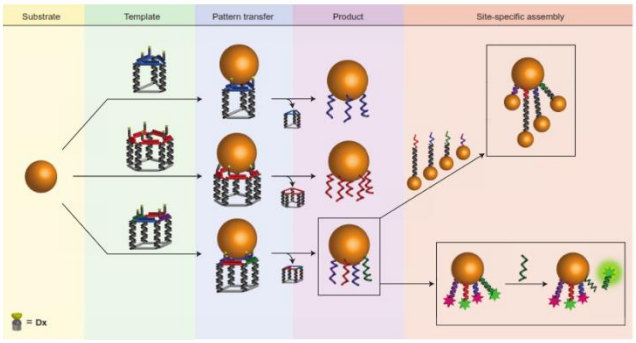


Figure 1.4: Patterning of oligomers using nanostructures[23]

with synthetic DNA, significantly reducing the material cost of these structures, albeit at the price of increased upfront expenses.

Having established the overarching objectives of biomimetics, and addressed some of the promising applications of DNA nanotechnology, we now move on to the chemical properties of biological DNA which make these feats possible.

Section 1.3: Relevant Properties of DNA

In biological systems, DNA stores genetic information which cells then transcribe into RNA, which is then transcribed into proteins, ultimately driving structure and behavior[6]. Genetic information is stored in the sequence of DNA polymer. All strands of DNA have the same backbone of alternating phosphate and sugar molecules. On the sugars of the backbone sit one of four nucleobases: Adenine (A), Thymine (T), Guanine (G), or Cytosine (C). The A and G bases have two conjoined rings and are chemical purines, while T and C

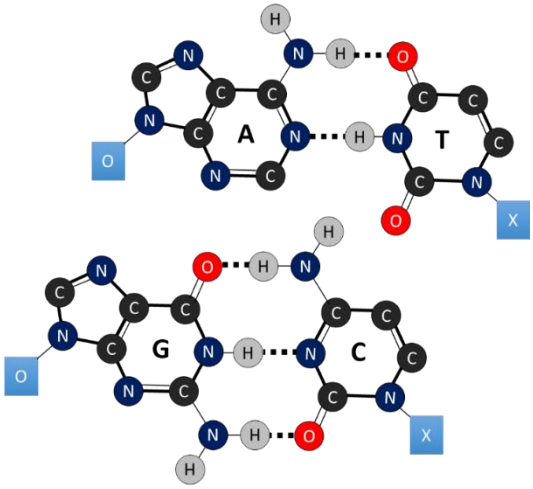


Figure 1.5: Watson-Crick base pairing, blue boxes indicate sugar/phosphate. O indicates 3'->5' direction entering the page, X indicates exiting the page

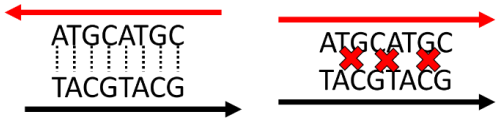


Figure 1.6: Watson-Crick base pairing and DNA directionality

have only one ring and are chemical pyrimidines. The cartoon in Figure 1.5 shows purines on the left and pyrimidines on the right, participating in Watson-Crick base pairing.

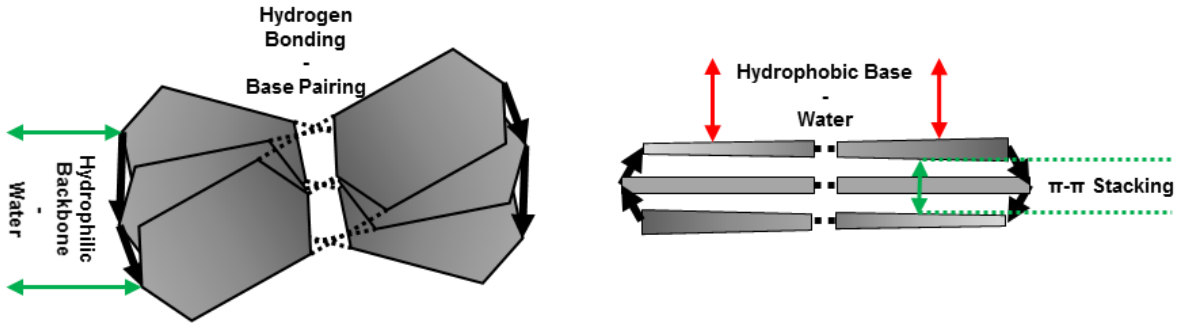


Figure 1.7: Hydrophobic and hydrophilic interactions in DNA self-assembly

Base paring results from the favorable hydrogen bonding between A and T bases or G and C bases. Because hydrogen bonding can only occur, for canonical systems, between these pairs, DNA can recognize regions of complementary sequence. It is this recognition between molecules that allows DNA to be read and replicated in the biological context.

The sugar and phosphate backbone of the DNA is chemically directional, as the phosphates are connected to the 3' and 5' carbons on the sugar molecule. While parallel strands can base pair by Hoogsteen base pairing, these are less common and less stable.

Despite the importance of base pairing, hydrogen bonding contributes surprisingly little to the thermodynamic driving forces in DNA assembly[6,29]. Figure 1.7 helps to illustrate why. Phosphate and sugar molecules are hydrophilic, while the nucleobases are

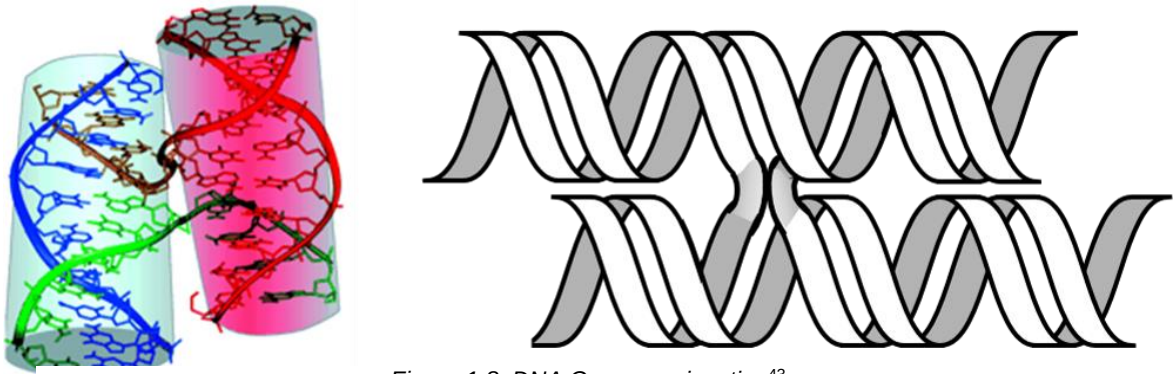


Figure 1.8: DNA Crossover junction[43]

hydrophobic conjugated rings. These hydrophobic and hydrophilic interactions provide the bulk of the energy driving DNA assembly, particularly with the favorable π-π stacking which occurs when double-strand DNA (dsDNA) forms. By forming dsDNA, the unfavorable base-water interactions are minimized to only those bases facing out into solution.

These properties can be utilized by designing sequences so that sequence-matching regions jump across multiple strands of DNA, as shown in Figure 1.8. Such jumps are not necessarily limited to conditions where unpaired bases float in solution. The crossover junction, depicted in Figure 1.8, allows a strand to jump between dsDNA helices without any interruption. These junctions can only occur when the rotating backbones of the two strands occur in the same location.

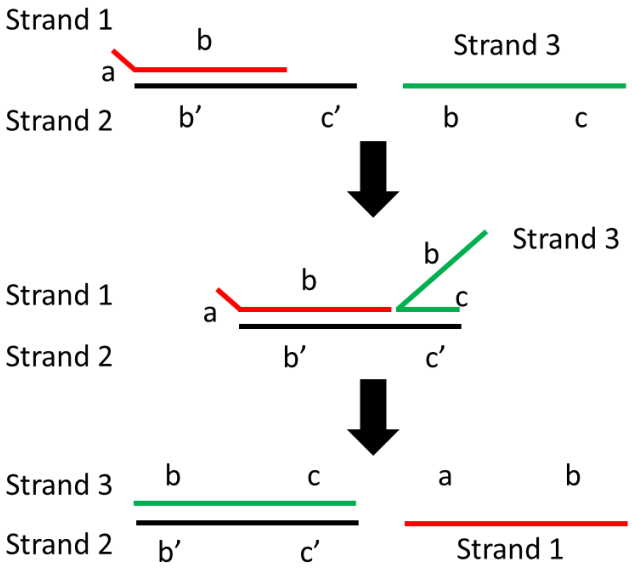


Figure 1.9: Strand displacement in DNA systems

A final relevant property of DNA is strand displacement. As a function of temperature, dsDNA is not perfectly bound together, as the hydrogen bonding itself is not the most powerful driving force. As such, strands of DNA can compete and exchange. An example of this is shown in Figure 1.9, which depicts three strands that compete for binding.

The sequences on each strand are labeled such that section n and n' are complements. In this example, both strands 1 and 3 contain the b sequence, and compete for the b' sequence on strand 2. The c' section is an unsatisfied toehold which is used by strand #3 to displace strand #1. These reactions are often used to perform computation and to drive switching in DNA nanostructures[30].

In this section we have addressed the most relevant chemical features of DNA: the general driving forces for DNA, its ability to bind together multiple helices, and its ability to perform reversible displacement reactions. In section 1.4 we will address the assembly of some DNA nanostructures and the driving forces particular to these systems.

Sections 1.4: DNA Nanostructures, Driving Forces, and Cooperativity

The driving forces for dsDNA formation from two ssDNA molecules have been well studied and characterized[6,29]. The nearest neighbor model allows one to rapidly convert the sequence of a dsDNA strand into a relatively accurate melt/annealing temperature. The model is based on the interactions of each base pair to the pairs above and below, as well as reductions in entropy associated with a transition from floppy ssDNA to the more rigid dsDNA[29]. Many aspects of the nearest neighbor model are important for DNA as a nanomaterial; chief among them is that GC bases are more energetically favorable than AT bases. However, the nearest neighbor model as implemented in common tools, like oligocalc or the IDT DNA calculator, provides insufficient detail to predict the properties of a nanostructure system[31,32].

To explain the differences between the thermodynamics of biological and of structural DNA as a nanomaterial, first we discuss scaffolded nanostructures in more detail. Figure 1.10 is a schematic for a small scaffolded nanostructure consisting of a scaffold and six synthetic staples. These staples only bind to the positions shown in the folded origami,

because their nucleobase sequences only match those regions. Figure 1.10 also shows a circle map of the same origami. The circle map shows the binding pads of each staple, with arcs between these pads. The arc length does not represent the staple; rather, it represents how far apart the binding pads are along the scaffold.

Surprising properties emerge from these systems due to a several extra-energetic interactions they exhibit, and the fact that there are many synthetic staples that interact with each other and guide the nanostructure formation. One such interaction that we have already discussed is the crossover junction; the energetics of crossover junctions is such that it has a negative enthalpy and a positive entropy[33]. The positive entropy is likely because the divalent cations at the junction can displace multiple water molecules bound to the phosphate backbone. For this reason, the stability of crossover junctions slightly improves at higher temperature. Another contributor to the energetics of DNA nanostructures is base stacking. If two staple ends touch each other, each receives a favorable change in energy due to the hydrophobic-hydrophobic interactions of the bases.

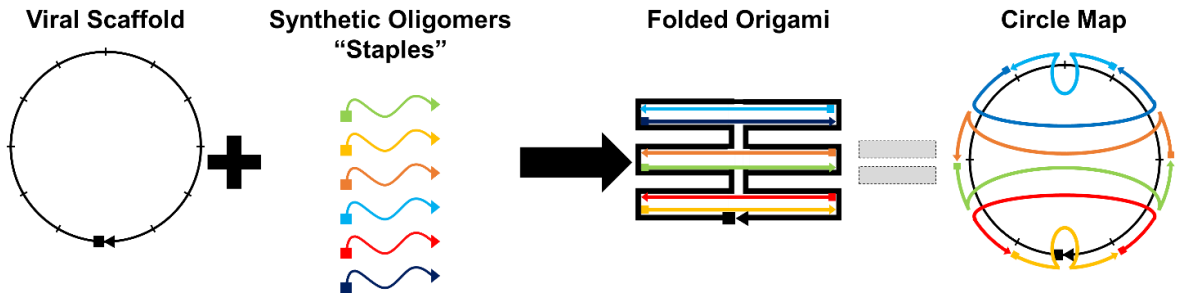


Figure 1.10: Scaffolded DNA nanostructures, or origami, routing map and circle map

Finally, and most directly relevant to the folding pathways of DNA origami, is the conformational entropy associated with a polymer loop. The entropy of a line becoming a loop, as a Jacobson-Stockmayer entropy extrapolation, as described by SantaLucia, is a good match with what one would predict from a polymer interpretation of statistical entropy,

and is shown in Eq. 1.1. In Eq. 1.1, C is a constant, R is the gas constant, T is the temperature in kelvin, N is the number of monomers in the loop, X is the number of monomers in a reference loop, and ΔS_x is the change in entropy for a smaller reference loop. The use of a reference loop improves the accuracy of this equation for smaller loops[29].

1.1) $\Delta S_{Line\rightarrow Loop} = -CRT \ln\left(\frac{N}{X}\right) - \Delta S_{Loop\ Length\ X}$

When modifying a loop, such as in scaffold/staple binding, the type of modification is vital. Figure 1.11 describes the two ways in which a staple could modify the scaffold loop. In the first path, the first staple merely reduces the size of the loop, while in the second path the first staple divides the scaffold into two smaller loops.

The change in entropy associated with the first staple binding for each of these paths is derived from Eq. 1.1 and shown in Eq. 1.2 and 1.3. The entropy change of the first path is dependent only on how much of the original loop was consumed. However, the entropy change of the second path depends strongly on the size of the two resulting loops. This second path has a much greater reduction in entropy than the former, most significantly when the size of the two resulting loops is equal.

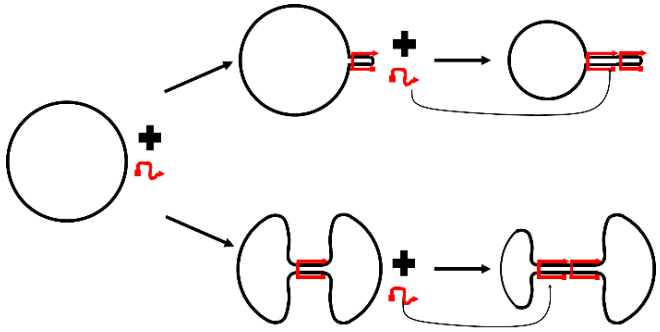


Figure 1.11: Configurational entropy reductions of staple binding to circular DNA

1.2) $\Delta S_{L1\rightarrow L2} = \Delta S_{L2} - \Delta S_{L1} = -CR \ln\left(\frac{N_1}{N_2}\right)$

1.3) $\Delta S_{Pinch} = (\Delta S_{L2} + \Delta S_{L3}) - \Delta S_{L1} - \Delta S_{DOF} = -CR \ln\left(\frac{N_2 N_3}{X N_1}\right) - \left(\frac{3}{2} + \frac{3}{2}\right) R - \Delta S_{loop\ length\ X}$

$$\Delta S_{Pinch} = -R\left(C \ln\left(\frac{N_2(N_1 - N_2)}{X N_1}\right) - 3\right)$$

For Eq. 1.3, there are several points of note. The first is the ΔS_{DOF}, or the change in entropy associated with the loss of degrees of freedom in the second loop as it is bound to the first. This change is 1/2RT for each of three translational degrees of freedom and three rotational degrees of freedom, as one would anticipate for a non-linear molecule via the equipartition theorem[34].

For this work the exact calculation of entropy reduction is not necessary. However, it is important to understand that the conformational entropy change penalizes the division of loops, and does so most strongly for cutting loops in half.

An important consequence of these forces is the rise of cooperativity in DNA nanostructure formation. This has been observed consistently in DNA nanostructure systems, and can be simply described as the tendency of any one strand to stabilize any and all neighboring strands. This leads to rapid assembly and disassembly events during annealing and melting respectively. Manipulating the cooperativity of DNA nanostructure systems could potentially improve both yield and control of folding, but first it must be more clearly understood.

In this section we discussed some of the theoretical nuances of these systems, particularly in crossover junctions, base stacking and conformational entropy. In section 1.5 we will discuss yield and processing of DNA nanostructures.

Section 1.5: Yield and Processing

Finding ways to leverage these unique tools was addressed briefly in section 1.4, but such applications are only half of the story. In order for any tool to be implemented, there must also be established techniques for prototyping and for optimizing yield at an industrial level. Much work has been directed towards this end, although more is still needed.

Yield has, for obvious reasons, received the bulk of our collective research. Techniques have been developed to improve annealing rates[35], staple motifs[36], and buffer conditions[37]. Other work has been done to quantify yield from low-cost methods such as gel electrophoresis.

Processing conditions have also been aggressively pursued. Much of the work addressing annealing rates has emphasized protocols that are compatible with other biomaterials[38,39]. Processing in non-aqueous solvent conditions[25] and in ionic liquids[40,41] has also been addressed. Still other work has dealt with the freeze-drying and storage of DNA nanostructures[26]. More notable work includes protocol modification to place DNA nanostructures on surfaces that must be protected from aqueous conditions[42].

While a truly robust theory of yield and processing is not yet complete, DNA nanotechnology has made strides towards this goal. In order to enable implementation at an industrial scale, we must also understand how the design of nanostructures affects yield and processing. This is necessary because the folding of any structure will be design-dependent, and the yield will depend on the folding. In section 1.6 we will discuss the general design of scaffolded structures.

Section 1.6: Design of Scaffolded DNA Nanostructures

As discussed in section 1.1 and 1.2, one of the strongest benefits of DNA nanotechnology is our clear understanding of the most basic and robust rules of assembly Watson-Crick base pairing. These rules provide the framework we need to design arbitrary shapes and structures, and an arbitrary number of ways to design a single structure. The plethora of design choices presents a daunting task for prototyping and experimental planning. Without a strong understanding of the relation between design and properties, it is difficult to make informed choices regarding structure design.

As shown in Table 1.1, the cost of design time contributes significantly to the up-front costs associated with DNA nanotechnology. The topic of design, even when limited to scaffolded nanostructures, contains incredible depth of which we have only scratched the surface.

Soon after the introduction of scaffolded DNA nanostructures in 2005, studies of nanostructure design[36]. These were, given limited understanding of the folding processes at the time, exhaustive. Since then, our understanding of the formation of DNA nanostructures has improved along with our characterization capabilities. With these new capabilities, we hope to better understand the relations between design and the folding of DNA nanostructures.

Figure 1.12 illustrates three areas in which choices must be made to design any DNA nanostructure. The first is typically the choice of a lattice, two of which are shown in Figure 1.12 and are viewed by looking down the long direction of the helices. Lattices are a

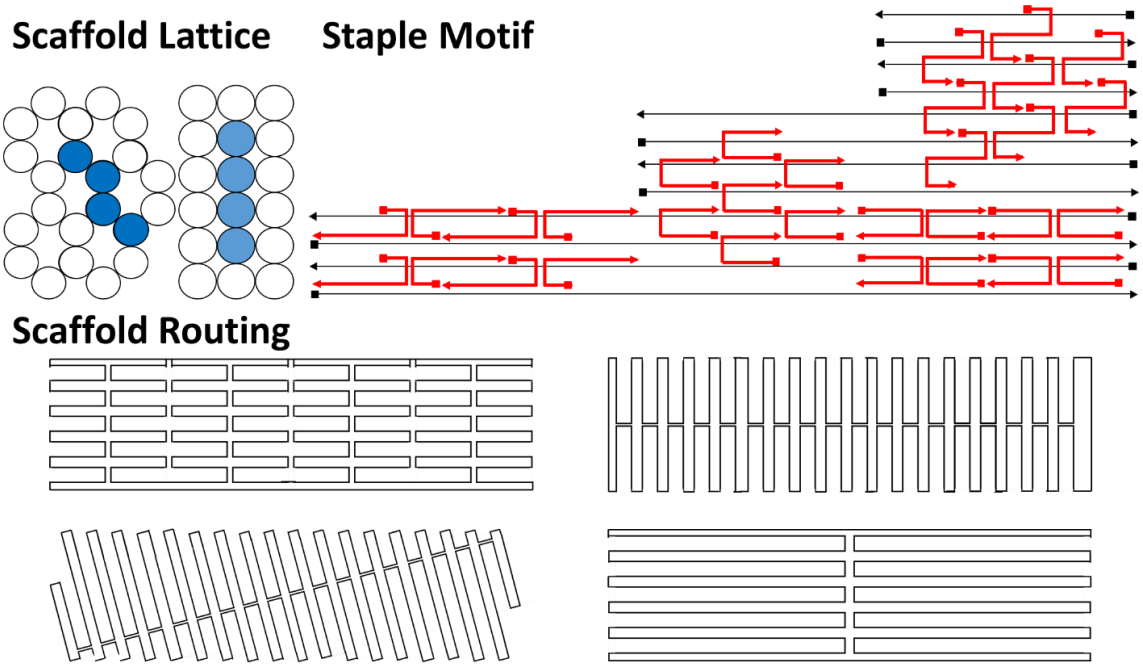


Figure 1.12: Design choices for scaffolded DNA nanostructures

way to describe the positions at which the helices cross over to one another. As shown in Fig 1.12, crossovers can only occur between helices when the backbones of the crossing strands meet. On any lattice, crossovers may occur between the same two helices at 32 base pair spacings. However, opportunities to progress down helices on a 2D plane occur every 16 base pairs on a square lattice, and on a hexagonal lattice, where helices are positioned at 60° angles, crossovers would occur at pairs of 10.6 and 11.2 base pairs of spacing. The effects of different lattices on stability and folding are not well understood.

Another necessary design choice is that of staple motifs, which are defined as the pattern of scaffold binding sections and crossovers which typically repeat for a given structure. These motifs are constrained by the base pair spacings at which crossovers are possible, but still provide many options. The motifs shown in Figure 1.12 are exclusively for a 2D origami on a square lattice. Much more work is necessary to understand the role these motifs play in structural stability, folding pathways, and contribution to the free energy of transition.

Finally, to create a structure one must choose a scaffold routing map. For a given 2D or 3D shape there are a large number of possible ways to route the viral ssDNA scaffold to form a structure. This is another area in which full understanding of design implications is lacking.

When considering these design choices in the light of section 1.5, one can imagine identical shapes with different designs and radically different energies of transition for individual staple binding events. This is particularly true for conformational entropy contributions, which should depend directly on the routing of a particular shape.

Having discussed, in general, our current state of understanding as regard to design and property relations and the plethora of uninformed design choices researchers must currently make, we move to the focus and scope of this work.

Section 1.7: Scope and Objective of this Work

Given the scope of our collective goals (section 1.1); the potential of DNA nanotechnology to achieve those goals (section 1.2); the chemical properties of DNA from which that potential stems (section 1.3); the unique and not perfectly understood properties of DNA, which become more relevant in nanotechnology than in biological systems (section 1.4); the processing and yield of DNA nanotechnology (section 1.5); and our understanding, or lack thereof, of the design and property relations which stem from these chemical properties and processing (section 1.6), we have established the necessity of further probing design and property relations as mediated by nanostructure folding.

As previous studies of design and property relations have reached their full potential due to the lack of appropriate tools, this work focuses on developing the toolkit for characterizing the formation and folding of DNA nanostructures.

In chapter 2, we will discuss the miniM13 workbench scaffold, whose high yield and smaller length significantly decreases both up-front and material costs for scaffolded DNA nanostructure systems. This reduction in cost is enables the rapid production of multiple designs, which may then be used in conjunction with the tools discussed in chapters 3-5.

In chapter 3, we will discuss gathering melt and anneal curves, which describe θ, the fraction of a system which is fully formed, as a function of temperature. Addressing the lack of baseline correction protocols, we provide an ensemble-level tool that improves our ability to quickly probe nanostructure formation with simple, low-cost experiments.

In chapter 4, we will introduce the technique of competitive multishape anneals in which two sets of staple pools compete for the same scaffold during annealing. This technique leverages the faster, less expensive prototyping of the miniM13, to probe which sections of a given structure succeed in competition. This in turn sheds light onto what, in the design of those structures, makes different regions more favorable.

In chapter 5, we will discuss another new technique to probe the folding of DNA nanostructures, that of chemical quenching. By addition of a quenching agent during annealing, we freeze DNA nanostructures mid-fold, allowing researchers to directly probe the folding pathways of DNA nanostructures.

In chapter 6, we will discuss how the synergy between these methods may be used to improve our understanding of design and property relations and ultimately the yield and quality of DNA nanostructures. We will do so through the lens of the model systems to which we applied our new techniques, the results of these tests, and how the combination of these results provides significant new insight into structure formation.

References:

1. Hilbert, M. & Lopez, P. 'The World's Technological Capacity to Store, Communicate, and Compute Information". *Science (80-.).* **332,** (2011).

2. UT-Battelle. Productive Nanosystems A Technology Roadmap. 198 (2007). doi:10.1088/0031-9120/40/4/003

3. Roco, M. & Bainbridge, W. S. Societal Implication of Nanoscience and Nanotechnology. *Technology* 272 (2001).

4. Vincent, J. F. V, Bogatyreva, O. A., Bogatyrev, N. R., Bowyer, A. & Pahl, A.-K. Biomimetics: its practice and theory. *J. R. Soc. Interface* **3,** 471–82 (2006).

5. Zahid, M., Kim, B., Hussain, R., Amin, R. & Park, S. H. DNA nanotechnology: a future perspective. *Nanoscale Res. Lett.* **8,** 119 (2013).

6. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. *Nucleic Acids: Structures Properties, and Functions*. (University Science Books, 2000).

7. Carlson, R. The changing economics of DNA synthesis. *Nat. Biotechnol.* **27,** 1091–4 (2009).

8. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).

9. Long, C. & Sweet, J. Quantitative selection of DNA aptamers through microfluidic selection and high-throughput sequencing. *South East Asia Res.* **14,** 445–469 (2006).

10. Zhang, D. Y. & Seelig, G. Dynamic DNA nanotechnology using strand- displacement reactions. *Nat. Publ. Gr.* **3,** 103–113 (2011).

11. Rangnekar, A., Nash, J. A., Goodfred, B., Yingling, Y. G. & LaBean, T. H. Design of Potent and Controllable Anticoagulants Using DNA Aptamers and Nanostructures. *Molecules* **21,** 202 (2016).

12. Hermanson, G. T. Bioconjugation techniques. *Acad. Press* **10,** 0123705010 (2008).

13. Samano, E. C. *et al.* Self-assembling DNA templates for programmed artificial biomineralization. *Soft Matter* **7,** 3240 (2011).

14. Douglas, S. M., Bachelet, I. & Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science (80-.).* **335,** 831–834 (2012).

15. Pilo-Pais, M., Goldberg, S., Samano, E., Labean, T. H. & Finkelstein, G. Connecting the nanodots: Programmable nanofabrication of fused metal shapes on DNA templates. *Nano Lett.* **11,** 3489–92 (2011).

16. Pilo-Pais, M., Watson, A., Demers, S., Labean, T. H. & Finkelstein, G. Surface-enhanced Raman scattering plasmonic enhancement using DNA origami-based complex metallic nanostructures. *Nano Lett.* **14,** 2099–2104 (2014).

17. Kuzuya, A. & Ohya, Y. Nanomechanical molecular devices made of DNA origami. *Acc. Chem. Res.* **47,** 1742–1749 (2014).

18. Yan, H., Feng, L., LaBean, T. H. & Reif, J. H. Parallel molecular computations of pairwise exclusive-or (XOR) using DNA 'string tile' self-assembly. *J. Am. Chem. Soc.* **125,** 14246–7 (2003).

19. Mao, C., LaBean, T. H., Relf, J. H. & Seeman, N. C. Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. *Nature* **407,** 493–6 (2000).

20. Won, R. Optical storage: DNA multiplexing. *Nat Phot.* **7,** 584 (2013).

21. Zheng, J. *et al.* From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **461,** 74–77 (2009).

22. Dutta, P. K. *et al.* DNA-directed artificial light-harvesting antenna. *J. Am. Chem. Soc.* **133,** 11985–11993 (2011).

23. Edwardson, T. G. W., Lau, K. L., Bousmail, D., Serpell, C. J. & Sleiman, H. F. Transfer of molecular recognition information from DNA nanostructures to gold nanoparticles. *Nat. Chem.* **8,** 162–170 (2016).

24. Ko, S. H. *et al.* High-speed, high-purity separation of gold nanoparticle-DNA origami constructs using centrifugation. *Soft Matter* **00,** 1–9 (2014).

25. Kim, H., Surwade, S. P., Powell, A., Donnell, C. O. & Liu, H. Stability of DNA Origami Nanostructure under Diverse Chemical Environments. (2014).

26. Rajendran, A., Endo, M. & Sugiyama, H. DNA Origami: Synthesis and Self-Assembly. *Curr. Protoc. Nucleic Acid Chem.* 3.1 (2000). doi:10.1002/0471142700

27. Ke, Y. *et al.* DNA brick crystals with prescribed depths. *Nat Chem* **6,** 994–1002 (2014).

28. Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485,** 623–6 (2012).

29. SantaLucia, J. & Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **33,** 415–440 (2004).

30. Zhou, L., Marras, A. E., Su, H. J. & Castro, C. E. DNA origami compliant

nanostructures with tunable mechanical properties. *ACS Nano* **8,** 27–34 (2014).

31. Kibbe, W. A. OligoCalc: An online oligonucleotide properties calculator. *Nucleic Acids Res.* **35,** (2007).

32. Owczarzy, R. *et al.* IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.* **36,** (2008).

33. Lu, M., Guo, Q., Marky, L. a, Seeman, N. C. & Kallenbach, N. R. Thermodynamics of DNA branching. *J. Mol. Biol.* **223,** 781–789 (1992).

34. Pathria, R. K. *Statistical Mechanics*. (Pergamon Press., 1972).

35. Sobczak, J.-P. J., Martin, T. G., Gerling, T. & Dietz, H. Rapid folding of DNA into nanoscale shapes at constant temperature.-supple. *Science* **338,** 1458–61 (2012).

36. Ke, Y., Bellot, G., Voigt, N. V., Fradkov, E. & Shih, W. M. Two design strategies for enhancement of multilayer–DNA-origami folding: underwinding for specific intercalator rescue and staple-break positioning. *Chem. Sci.* **3,** 2587 (2012).

37. Martin, T. G. & Dietz, H. Magnesium-free self-assembly of multi-layer DNA objects. *Nat. Commun.* **3,** 1103 (2012).

38. Sobczak, J.-P. J., Martin, T. G., Gerling, T. & Dietz, H. Rapid folding of DNA into nanoscale shapes at constant temperature. *Science* **338,** 1458–61 (2012).

39. Zhang, Z., Song, J., Besenbacher, F., Dong, M. & Gothelf, K. V. Self-assembly of DNA origami and single-stranded tile structures at room temperature. *Angew. Chemie - Int. Ed.* **52,** 9219–9223 (2013).

40. Gállego, I., Grover, M. a. & Hud, N. V. Folding and Imaging of DNA Nanostructures in Anhydrous and Hydrated Deep-Eutectic Solvents. *Angew. Chemie Int. Ed.* n/a–n/a (2015). doi:10.1002/anie.201412354

41. Zhao, H. DNA Stability in Ionic Liquids and Deep Eutectic Solvents. *J. Chem. Technol. Biotechnol.* n/a–n/a (2014). doi:10.1002/jctb.4511

42. Lee, J. *et al.* The restoration of DNA structures by the dry–wet method. *Soft Matter* **8,** 619 (2012).

43. Eichman, B. F., Vargason, J. M., Mooers, B. H. M. & Ho, P. S. The Holliday junction in an inverted repeat DNA sequence: Sequence effects on the structure of four-way junctions. *Proc. Natl. Acad. Sci.* **97,** 3971–3976 (2000).

Chapter 2: MiniM13 Workbench Scaffold

Section 2.1: Introduction & Motivation

In this chapter we will discuss the miniM13 scaffold, which we created as a high yield rapid prototyping/workbench scaffold. The miniM13 has a shorter scaffold length of 2,404 bases, compared to the standard M13 length of 7,450 bases. As such, it has 1/3 the upfront staple cost. This will improve the cost incentives for studies which address differing design motifs in scaffolded DNA nanostructures.

The design capabilities of DNA self-assembly are, for the purposes of nanofabrication, one of its primary appeals. For scaffolded DNA nanostructures, the primary design limitations are those of chemistry and geometry described in sections 1.3-1.6, along with the length of the scaffold strand. For this reason, many research groups have developed modified scaffolds, both longer and shorter than the standard M13[1–3]. As discussed in section 1.1, more than 80% of the cost of scaffolded nanostructures is the viral scaffold[4]. Low cost, high volume production of the viral scaffold DNA is possible, but lack of a broad market has thus far prevented this.

The scaffold itself plays a role in design through more than just cost and limiting maximum structure size. A large ssDNA strand connected to the nanostructure could interfere with its function; as such scaffolded structures typically use most or all of the available scaffold. The M13 scaffold can be bought in small volumes, 10 µg or 4.4 picoMoles, while the minimum production scale of synthetic DNA is approximately 80 nanoMoles. At 80% yield, this purchase size is equivalent to 44,000 anneals and ~$1,300 before the purchase of strands with conjugation chemistry. Since the number of synthetic staples is linked to scaffold length, the upfront costs are highly dependent on the staple length.

An additional cost of design is the use of personnel. In section 1.1 we briefly discussed the upfront and material costs of scaffolded structures, in which design costs were included. Here we will discuss design costs in greater detail. An estimate design/prototype turnaround for M13 scaffolded structures is between one and six months, averaging ~3 months. Design tasks are often undertaken in parallel with other duties, so we assume 4 hours per workday are spent on design. At a pay of $10 / hr, one can anticipate an upfront expense of $2,400 on average for an M13 structure.

Between design and minimum purchase sizes, implementation of a M13 scaffolded structures often costs ~$3,700. This upfront cost encourages researchers to build systems which can perform multiple experiments. Typically, this is done by switching out sets of labeled strands in a single structure[5–7]. However, doing so requires the structures be designed with potential locations for labels in mind, which requires additional time. After purchase and annealing, flaws in a system often take as much as a month to identify and correct. Altogether these factors encourage researchers to invest more person hours into designing DNA nanostructures. This creates a cycle of increased overall cost, and of fewer new structures being designed.

This design cost has been reduced in many ways since 2005. The introduction of design and predictive software have reduced the time expense and the number of prototype cycles associated with design. Protocol modifications minimize the use of staple excesses, and avoid expensive additional purification steps. Finally, the concept of a workbench scaffold reduces both the initial cost and design time by the available size of the nanostructures[1].

While protocols which include creation and use of smaller scaffolds have existed[2], to our knowledge the first such scaffold designed specifically for adoption outside of its home lab is the M1.3. The M1.3 workbench scaffold has been created by using ECor1 and Bg1II

restriction enzymes to cut a section from a standard M13 scaffold, followed by reforming the circle with T4 DNA ligase. The M1.3 is 704 bases long, and while not excessively difficult to produce, requires some degree of molecular biology experience. Additionally, this scaffold size results in structures that are difficult to image in AFM, requiring TEM or other techniques. A slightly larger and easier to synthesize scaffold would thus be of value to the DNA nanotechnology community.

This chapter discusses the concept of a workbench DNA scaffold which is low cost, easy to produce, and large enough to create AFM visible nanostructures while being small enough to have a manageable design/prototype cycle. To achieve this goal a collaboration was established between the LaBean lab at NC State and Dr. Stanley Brown at the University of Copenhagen; Dr. Brown created the scaffold production system which we implemented at NC State and used the scaffold to create a variety of nanostructures.

Section 2.2: The miniM13 Scaffold

Before describing the methods of producing the miniM13 scaffold, it is worth briefly providing background on bacteria and viruses. Unlike mammalian cells, bacteria do not have a nucleus; the DNA which bacteria read and use to create proteins float freely within the cell. Viruses like M13 consist of DNA and a protein shell. The protein shell, by various means for different bacteria, provides the DNA entrance into the bacteria. At a minimum, the virial DNA must contain sequences which create more of the protein shell and which begin replication of the viral DNA. Viruses often contain much more than this minimum. However, as mutations can randomly occur, there is an evolutionary pressure to weed out any sequence which does not actively contribute to replication of the virus[8,9].

Given that the viral DNA must encode all of this information, there are limits to how small one can make a fully functioning virus. The miniM13 system circumvents this through

the use of helper phages. Helper phages encode the protein shell with minimal production of viral DNA. As such, the sequences for protein production were removed from the miniM13 without reducing yield of virus. Helper phages do this as they have a defective origin of replication[10]. The origin of replication is a DNA sub-sequence at which a cell will begin copying the entire DNA strand. A defective origin minimizes the replication of the helper phage DNA while continuing to produce the packaging proteins which are encoded on other portions of the phage[3,11]. This strategy has been previously used to create a DNA mini-box[2].

Dr. Brown developed a helper phage which had a deleted, rather than defective, origin of replication as well as the sequence for chloramphenicol antiobiotic resistance[12]. The deleted origin of replication entirely prevented creation of unwanted viral DNA. The sequence for chloramphenicol resistance ensured that any bacteria without the helper phage would die, if the growth media was supplemented with chloramphenicol. This greatly improved production of packaging proteins.[11]

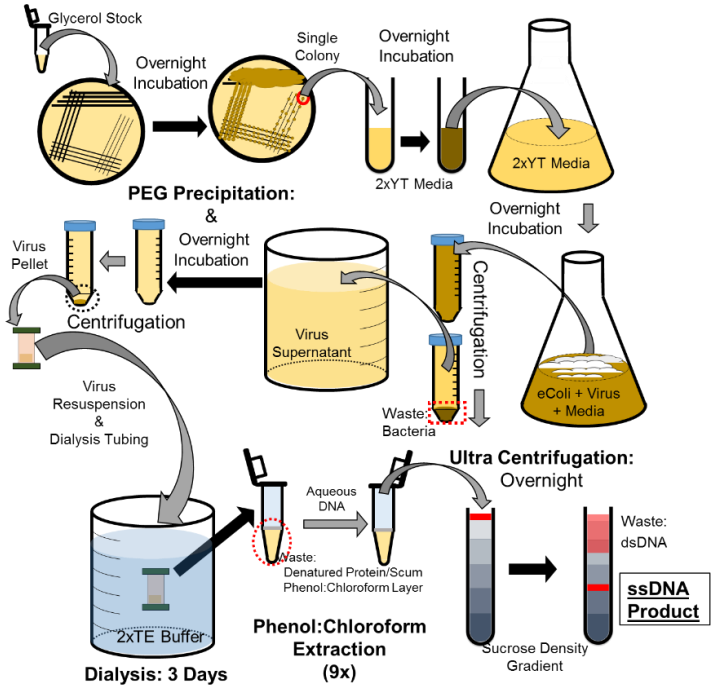


Figure 2.1. Schematic of miniM13 scaffold production

The miniM13 contains the M13 origin of replication, as well as the sequence for ampicillin resistance. In culture media containing both chloramphenicol and ampicillin, the only surviving eColi should then be those with both the helper and miniM13 phages. This improved yield both in production and in fewer purification steps.

Production protocols of the miniM13 scaffold were adopted with minimal changes from the Brown lab; Modifications were made to accommodate for available equipment and for culture volumes greater than 2 Liters. Figure 2.1 shows the general flow of cell culture for miniM13 production. The S3131 eColi strain containing both the miniM13 and helper phage were plated from glycerol stocks, from which single colonies were selected and grown in rich culture media with chloramphenicol and ampicillin. After growth, the bacteria and bacterial waste were centrifuged from solution. Polyethylene glycol (PEG) was then added to the supernatant, which precipitated the virus after overnight incubation. The virus was resuspended in a small volume and dialyzed for three days to further remove protein waste. The DNA was removed from the viral proteins using phenol:chloroform extraction; Phenol denatures proteins, forcing them to leave solution, while chloroform encourages phase segregation of the phenol and

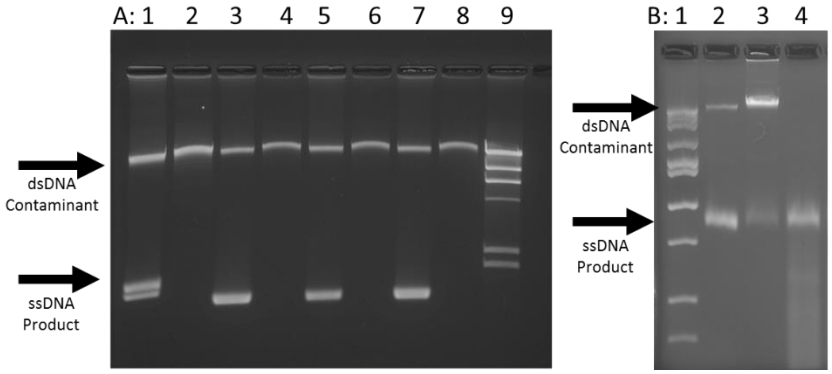


Figure 2.2: A-1.5% Agarose gel by Dr. Brown. Lanes 1-8 each contains DNA from 1 mL of culture. Lanes 1 and 2 are from eColi with the parent helper phage pBluescript. Lanes 3-8 are from independant cultures of S3131. Samples in odd lanes were subjected to S1 nuclease which attacks ssDNA. Lane 9 is lambda DNA digested with HindIII. B- 2% Agarose gel performed at NC state, Lane 1 is the Fisher Hi-Lo ladder, Lane 2 is the miniM13 from culture, and Lanes 2 and 3 are the waste dsDNA and product ssDNA after sucrose gradient centrifugation. Smearing of lane 4 suggests minor damage occurred during purification

proteins from the aqueous layer. After extraction the DNA is precipitated and waste DNA is removed by either isopicnic or sucrose gradient ultracentrifugation. S1 nuclease, which only attacks ssDNA, is used to confirm that the band migrating at 2,404 bases is ssDNA. Figure 2.2 A shows S1 nuclease assays performed in the Brown lab. Figure 2.2 B shows the unpurified DNA after phenol:chloroform extraction, and the sucrose gradient purified dsDNA contaminant and ssDNA product.

In this section we have described the theory behind the miniM13 system, outlined its production, and shown it's initial results. In section 2.3 we will address the yield of miniM13 production as compared to the M1.3 and the M13 scafolds.

Section 2.3: Yield and Cost

Dr. Brown obtained yield of the miniM13 of approximately 400 µg/L of culture media. We obtained approximately 360 µg/L in our implementation[10]. This compares favorably to the M1.3 workbench scaffolds which obtained approximately 0.1-1 µg / L[1]. Both protocols take approximately 4-5 days. Material costs for one production cycle of both miniM13 and M1.3 scaffolds are approximately the same, estimated at €22 or $24. The M13, as of 2016, costs approximately $37 / 10 µg. Given these yields Table 2.1 shows the relative upfront costs of the M13, miniM13, and M1.3 scaffolds.

Table 2.1: Estimated upfront costs of implementing new scaffolded nanostructures for the M13, M1.3, and miniM13 Scaffolds

Upfront Costs	***M13***	***miniM13***	***M1.3***
Scaffold Length	7,249 bp	2,404 bp	703 bp
Scaffold Cost	$3.7 / µg	$1.2 / µg	$420 / µg
(minimum production size)	10 µg	400 µg	1 µg
Upfront Scaffold Cost	$37	$420	$420
Upfront Staple Cost	$1,305	$433	$127
Est. Design Cost	$2,400	$795	$350
Est. Upfront Cost for New Nanostructures	**$3,742**	**$1,648**	**$907**

Table 2.1 shows the estimated cost of each scaffold. Person hours were included in scaffold cost and in design time assuming $10/hr pay and a 40 hr work week. Design time for miniM13 structures, discussed in section 2.4, was consistently ~1 month. As the average time for M13 structure design is 3 months design time for M1.3 was estimated to be ~9 days assuming a linear relationship between design time and scaffold length.

The impact of these scaffolds on materials cost is shown in Table 2.2. Anneals often have a similar mass of nanostructures to streamline imaging. As such, the miniM13 anneal was considered to be 15 nM and with M1.3 being 45 nM, compared M13 at 5 nM. Initial estimates in 2005 were corrected for inflation, for the use of 10x, rather than 100x, excess staple concentration and for the neglect of PAGE purification of staple oligomers. Since protocols for DNA annealing have been partially optimized in the last 10 years we included costs both before and after optimization. While anneal sizes vary, one nanoMole of scaffold will create >6,000 anneals, where each anneal is 30 µL at 5 nM. If the nanostructures are decorated and/or purified, the yield decrease necessitates much larger sample sizes.

Table 2.2: Estimated material costs for the M13, M1.3, miniM13 scaffold systems in nMoles and Anneals

Material Costs	***Scaffold / nMole***	***Staples / nMole***	***Material Cost / nMole***	***Material Cost*** */ anneals*
2005 M13 Estimate *Adjusted for inflation*	$9,322	$2,361	**$11,683**	**$1.75**
2005 M13 Estimate *w/ 10x Stap. No PAGE*	"	$244	**$9,566**	**$1.43**
Current M13 Estimate	"	$198	**$9,448**	**$1.42**
M1.3 Estimate	$91,107	$19	**$91,126**	**$136.70**
miniM13 Estimate	$865	$66	**$931**	**$0.42**

As shown in Table 2.2 and Table 2.1, while both the minimM13 and M1.3 have a significantly lower upfront cost than the M13 scaffold, the M1.3 is significantly more expensive per mole and per anneal than the M13 origami. This is due to the low yield of M1.3 production;

the reported M1.3 yield is at least 300x lower than miniM13. The miniM13 reduces upfront costs by half, and material costs to 30% of the M13 structures.

In this section we have outlined the yield of miniM13 production. The reductions in cost associated with the miniM13 should encourage experiments for which multiple designs are necessary. In the section 2.4 we will briefly discuss the physical properties of the miniM13 phage.

Section 2.4: Persistence Length

Because of its rod like structure, the M13 phage it has been used as a biological liquid crystal[13]. The structure of the miniM13 should be identical to the M13 save only in that it contains less DNA packaged in fewer protein molecules. As such, one would expect the intrinsic properties of the miniM13 to be identical to M13. As the persistence length is relevant

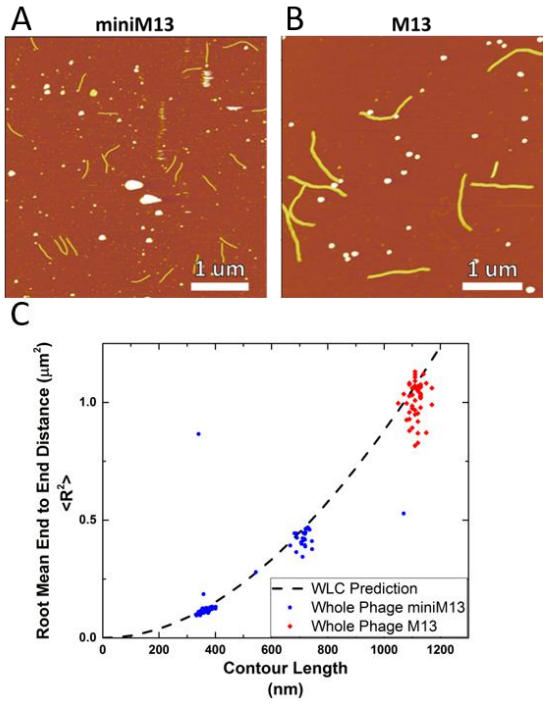


Figure 2.3: A- whole miniM13 phage, B- whole M13 phage, C- phage contour/end-to-end distances compared to WLC theory prediction

to the ability to form liquid crystals, we compared the persistence length of both the M13 and miniM1 via AFM imaging.

According to Worm Like Chain theory, or WLC, a 3D chain on binding to a 2D surface will have a root mean squared end to end distance, $\langle R^2 \rangle$, determined by Eq. 2.1[14], where persistence length is P and contour length is L[14–16].

2.1) $$\langle R^2 \rangle = 4PL\left[1 - \frac{2P}{L}\left(1 - e^{\frac{-L}{2P}}\right)\right]$$

Figure 2.3 shows example AFM images of the miniM13 and M13 phage, as well as a plot of their $\langle R^2 \rangle$ and contour length. Data from both whole phage samples fell along the prediction obtained from WLC theory using Eq. 1 and the literature value for M13 persistence length, 1,265 nm. A minority of the phage in the miniM13 sample had twice the length of the other phage. Given the dilution of the sample and the width of the phage, it is unlikely that these were two stacked miniM13 phage or a dimer DNA molecule. Rather, it appears that two miniM13 molecules were packaged in the same protein sheath. This is confirmed by the lack of a dimer ssDNA band in gel electrophoresis of miniM13 systems such as Figure 2.2.

In this section we have briefly discussed our characterization of the miniM13 phage persistence length, and confirmed its consistency with M13 and WLC theory.

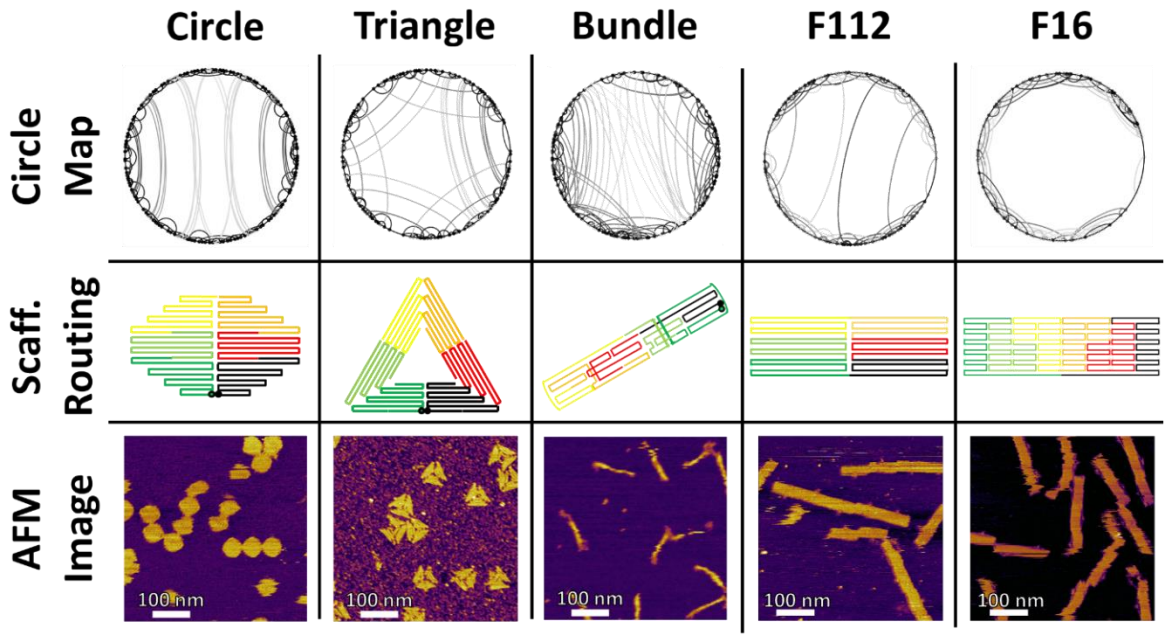
Section 2.5: Nanostructure Design with miniM13 Scaffold

Design of miniM13 DNA nanostructures is identical to that of M13 structures. First the footprint of the structure is roughly sketched, and the approximate maximum size of the structure is calculated. As per section 2.1, this is due to the potential issues associated with too much excess ssDNA scaffold hanging off of the final structure. Typically, the routing is selected by the most symmetric/simple scaffold routing. Since the effects of scaffold routing are not entirely known, this is partially for aesthetic purposes and partially to keep the design

process straightforward. Undergraduate researchers were trained by members of the LaBean lab to design multiple origami.

Table 2.3 shows five of the six origami designed thus far using the miniM13 scaffold. The top row shows the distances across the scaffold which the staples bind, the second shows the scaffold routing, and the bottom row shows liquid tapping mode AFM of the origami. The 6th origami, the I-motif origami, identical to the F112 except in that it has 6 seam positions rather than five, and they are bridged by actuating I-motif sequences. Of the 6 origami four were designed by undergraduate researchers, namely the circle, bundle, triangle and I-motif. We designed the F16 and F112 as a model system to be discussed in chapter 5. All of these structures were designed in less than a month of person hours.

Table 2.3: miniM13 Nanostructures- circle maps, scaffold routing, and AFM images



In this section we have discussed structures which we have designed for the miniM13 scaffold and the relative ease with which the design was performed. In section 2.6, we will address future work for the miniM13 project.

Section 2.6: Future Work

The miniM13 workbench scaffold reduces both the upfront and material costs for newly designed systems by approximately half and material costs by 70%. These cost reductions open significant opportunities to perform experiments in which multiple new structures are designed. In particular, experiments which examine design/property relations will benefit from the miniM13 scaffold.

There are two significant thrusts for the future of the miniM13. The first is commercial production of the scaffold, which will further reduce costs and production time. The second is to probe the tendency for multiple miniM13 molecules to be packed into a single phage packet, as seen in Fig 3 A&C. While this tendency does not appear to influence yield, it is an interesting feature which bears investigation.

References:

1. Richert, C. *et al.* M1.3 - A Small Scaffold for DNA Origami. *Nanoscale* 284–290 (2012). doi:10.1039/c2nr32393a

2. Andersen, E. S. *et al.* Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459,** 73–76 (2009).

3. Marchi, a. N., Saaem, I., Vogen, B. N., Brown, S. & LaBean, T. H. Towards larger DNA origami. *Submitted* (2014). doi:10.1021/nl502626s

4. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).

5. Samano, E. C. *et al.* Self-assembling DNA templates for programmed artificial biomineralization. *Soft Matter* **7,** 3240 (2011).

6. Castro, C. E. *et al.* A primer to scaffolded DNA origami. **8,** 221–229 (2011).

7. Liu, Q., Song, C., Wang, Z. G., Li, N. & Ding, B. Precise organization of metal nanoparticles on DNA origami template. *Methods* **67,** 205–214 (2014).

8. Lambert, P. a. *Introductory microbiology*. *Trends in Microbiology* **3,** (1995).

9. Sambrook, J. & Russell, D. W. *Molecular Cloning - Sambrook & Russel*. *Human Mutation* **18,** (2001).

10. Short, J. M. *et al.* Volume 16 Number 15 1988 Nucleic Acids Research. *Nucleic Acids Res.* **16,** 7583–7600 (1988).

11. Brown, S., Majikes, J., Fennell, H., Samano, E. C. & Labean, T. H. An easy-to-prepare mini-scaffold for DNA origami. *Nanoscale* **4434,** 16621–16624 (2015).

12. Brown, S. *et al.* An Easy-to-Prepare Mini-Scaffold for DNA Origami. *Nanoscale* (2015). doi:10.1039/C5NR04921K

13. Chung, W.-J., Merzlyak, A. & Lee, S.-W. Fabrication of engineered M13 bacteriophages into liquid crystalline films and fibers for directional growth and encapsulation of fibroblasts. *Soft Matter* **6,** 4454 (2010).

14. Abels, J. a, Moreno-Herrero, F., van der Heijden, T., Dekker, C. & Dekker, N. H. Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys. J.* **88,** 2737–2744 (2005).

15. Khalil, A. S. *et al.* Single M13 bacteriophage tethering and stretching. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 4892–4897 (2007).

16. Brinkers, S., Dietrich, H. R. C., Groote, F. H. De & Young, I. T. The persistence length of double stranded DNA determined using dark field tethered particle motion. 1–9 (2009). doi:10.1063/1.3142699

Chapter 3: DNA Nanostructure Anneal/Melt Curves via Fluorescence Spectroscopy

Section 3.1: Introduction and Motivation

For material processing, understanding reversible phase changes is crucial. Reliable denaturing/hybridization data is essential. However, for DNA nanostructures there are experimental barriers to obtaining such information. The sensitivity of measurement techniques particularly at typical concentrations for DNA nanostructures and the material cost of performing such measurements are both barriers. Obtaining the fraction of a DNA nanostructure system in the ssDNA or dsDNA state is surprisingly non-trivial, particularly as it is an ensemble level measure.

As discussed in section 1.4, the thermodynamics of biological DNA have been thoroughly studied and understood[1,2]. While most, if not all, of the individual theoretical components for predictive understanding are relatively well known, such as energetics of double crossover formation[3], they are only just being incorporated into usable models for complex, artificial nanostructures[4,5]. When transitioning from established systems such as tiles to larger structures such as scaffolded DNA origami, multiple technical difficulties arise. Two sources of these difficulties are the overall cost of such systems and the excess ssDNA used to drive DNA nanostructure formation. The overall cost becomes nearly prohibitive for many of the less sensitive techniques, while the excess ssDNA disrupts the signal to noise ratios in others.

Much of the work on processing DNA structures has avoided this pitfall by the use of final assembly yield as a metric, as measured by AFM. However, as our understanding of processing improves, this gap in capabilities becomes more problematic[6,7]. Several attempts have been made to bridge this gap. The Yan group has put individual fluorophore/quencher pairs within a scaffolded nanostructure system to determine localized melting[8]. This provides

detailed information for discussions of cooperativity, as defined in section 1.4, but is prohibitively expensive for general use. To our knowledge only the Elezgaray lab has successfully performed UV-VIS spectroscopy with publication quality data for melting on DNA nanostructures, to confirm their predictive program[9,10].

Finally, while in pursuit of other objectives, three groups have performed fluorescence spectroscopy, using the SYBR Green I dye, with the derivative of fluorescence rather than the fraction of dsDNA, θ, as a figure of merit[4,5,11,12]. Obtaining θ requires baseline corrections, while the derivative can be performed on raw data. The Turberfield/Ouldridge groups used the derivative of fluorescence data to verify formation prediction models[4,11]. The Dietz and Estevez-Torres groups used the derivative of fluorescence data to inform new experimental techniques[5,12]. While dθ/dT was sufficient for the goals of those studies, the absence of baseline correction is notable, as such corrections are common for biological DNA systems[13,14].

Reliable calculation of θ is desirable for more than the general purposes discussed above. Knowing how much of a structure has formed as a function of temperature and processing conditions will provide a valuable stepping stone between those conditions and final yield. Further, comparison of melting and annealing data would allow quantification of properties such as hysteresis and temperature range of the transition which are both measures of cooperatively. This information will link otherwise equivalent design choices for DNA nanostructures with their effect on thermal properties and yield. Design of nanostructures on an industrial scale will demand these capabilities.

In this section, we have discussed the motivation, and background, for obtaining ensemble level information regarding the fraction of dsDNA formation, or θ, of DNA

nanostructure systems. In the next section, we will briefly address idealized annealing and melting data for such systems.

Section 3.2: Annealing and Melting Data

The fraction of a system which has formed dsDNA, or θ, for a short dsDNA strand takes the form of a sigmoidal curve representing single reversible reaction. The sigmoidal is centered on the transition temperature, T_M/T_A, which is defined as the temperature at which θ equals 0.5, shown in Figure 3.1 A. The θ can be described by the Arrhenius equation for a reversible reaction at equilibrium. For systems with longer DNA strands or multiple dsDNA reactions, this transition can become significantly more complicated as in Figure 3.1 D. The hysteresis between melt and anneal curves, is measured by the integral between them, Figure 3.1 B. As the θ curve is defined by change in Gibbs free energy of the transition, the hysteresis

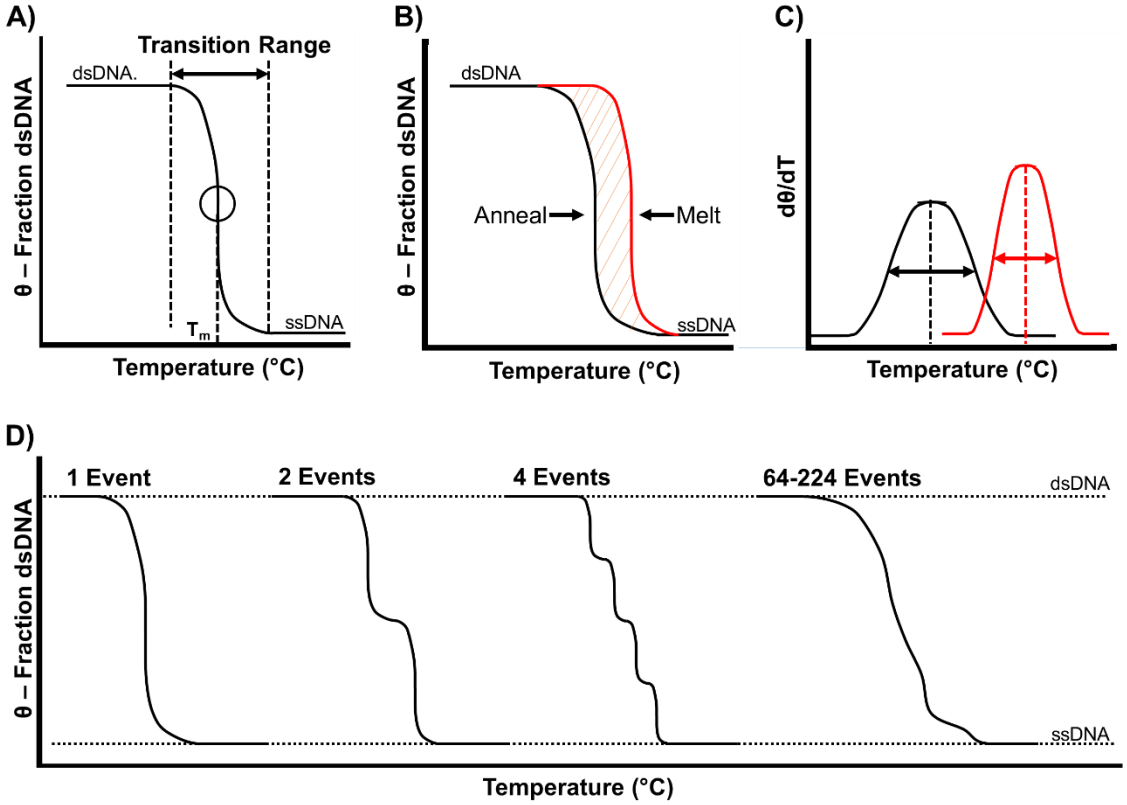


Figure 3.1: Idealized curves for melt/anneal data. A-a typical melt curve with transition breadth and Tm marked, B-hysteresis area between melt and anneal, C-dθ/dT, D-Change in θ curves for complex systems

is a measure of path dependent effects on ΔG, specifically cooperativity for scaffolded DNA nanostructures. For many purposes, the derivative of θ as a function of temperature is used rather than θ itself as shown in Fig. 3.1 C; this minimizes the need for baseline corrections, and provides a reasonable approximation of the transition temperature and transition range for single step reactions[15].

For scaffolded DNA nanostructures, the ensemble θ information becomes significantly more powerful. As illustrated in Figure 3.1 D, the transition between dsDNA and ssDNA is significantly broader for nanostructures. However, the cooperativity of individual binding events shortens this transition as each binding/unbinding is made more favorable by the preceding event. As such, θ contains two measures of cooperativity in its transition breadth and hysteresis. Some of this information can be implied by the dθ/dT curves, particularly in the narrowing of the melt. However, θ itself provides a much more reliable measure.

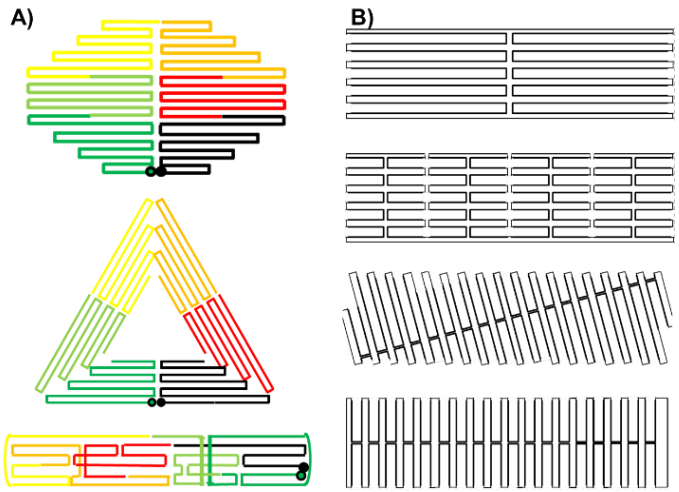


Figure 3.2: Example routing patterns for comparison. A- distinct DNA nanostructures, B- structures of identical shape but different design

Additionally, θ as function of temperature provides the crucial ability to compare the formation of different nanostructures. These comparisons are valuable for comparing distinct structures,

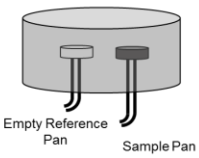
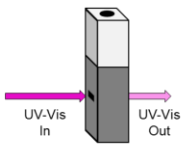
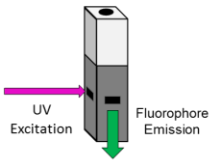
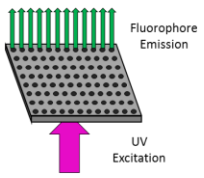
colored routing maps in Figure 3.2 A, as well as for testing how design choices such as scaffold routing can affect formation, Figure 3.2 B.

In this section we have discussed idealized melting and annealing data for DNA systems, and how such information would be useful within the context of DNA nanostructures. In section 3.3, we will discuss common techniques used to gather such information, as well as the advantages and disadvantages of these techniques.

Section 3.3: Common Experimental Techniques

One approach to experimentally probing reversible processes is to use techniques designed specifically for thermal properties such as Differential Scanning Calorimetry, which measure the phase transitions and thermal properties. Alternatively, one can modify a characterization technique often used on that material, for example X-Ray diffraction or Atomic Force Microscopy, to perform across a broad range of temperatures. The latter is discussed in detail in Chapter 4. Table 3.1 shows four techniques in the former category which could be used to analyze a DNA nanostructure system. Differential Scanning Calorimetry delivers by far the most direct measurements of thermal properties; in a DSC the amount of additional energy required to heat/cool between a sample and reference is monitored.

Table 3.1: Common experimental techniques for anneal/melt data[30–32]

	Differential Scanning Calorimetry	**UV-Vis Spectroscopy**	**Fluorescence Spectroscopy**	**Fluorescence: RT-PCR Equipment**
	Thermally Insulated Chamber 	Quartz Cuvette 	Quartz Cuvette 	96 Well Plate 
Measurement	Heat Flow heat capacity (C_p)	Optical absorbance λ = 260 nm	Fluorescence reporter molecule	Fluorescence reporter molecule
Sample Size & Throughput	~0.5 mg/mL~500 µL 1 sample/run	~400 nM ~200 µL <6 samples/run	~15 nM ~200 µL <3 samples/run	<15 nM ~20 µL <96 samples/run

However, given the small amounts of energy change associated with DNA melting, DSC requires significant volumes of high concentration DNA[3]. Even high resolution nano-DSCs designed for biomolecules require concentrations of approximately 0.5 mg /mL. This expense is prohibitive.

Ultraviolet-Visible spectroscopy, or UV-Vis, probes the changes in absorption associated with nucleobase motion as a function of temperature[15]. These changes in absorption track roughly with the reduction in motion associated with base pairing, base stacking, and the ssDNA/dsDNA transition[1]. Figure 3. A and C show the raw and processed data for an idealized UV-Vis melting experiment. UV-Vis requires much lower concentrations of DNA than DSC, making it an economically possible choice for DNA-based nanostructures. However, the excess ssDNA added to drive annealing, particularly in scaffolded origami systems, significantly lowers signal/noise ratios for UV-VIS. Increasing concentration can improve the signal, but any given spectrophotometer is limited in how much absorption can be measured. As such, both DSC and UV-VIS have significant limitations to the concentrations which they can reasonably measure, both are unable to address typically used sample concentrations, and neither can be considered inexpensive for DNA nanostructures.

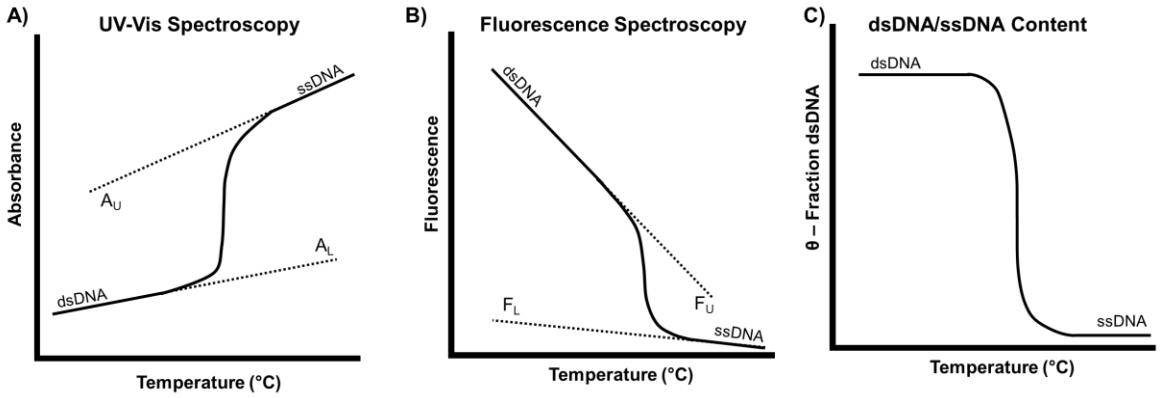


Figure 3.3: Experimental data before and after processing A- UV-VIs, B-Fluorescence, C-final θ curve

Another common method for determining ssDNA/dsDNA concentration is fluorescence spectroscopy. A fluorescence spectrophotometer is typically set up such that the light beam travels through a cuvette exciting fluorophores and, from a perpendicular direction, the fluorescence emission spectrum is measured as shown in Table 3.1[16]. Fluorescence spectroscopy systems are similar in design to UV-Vis systems and often include UV-Vis capabilities as both excitation and emission occur in the UV-Vis spectra range. Fluorescence spectroscopy requires some set of probe molecules to report the ssDNA/dsDNA state of the DNA.

One probe molecule strategy is that of a fluorophore/quencher pair covalently bonded to the DNA. Fluorophore/quenchers are Förster resonance energy transfer donors and acceptors respectively[17,18]. When the pair are closely spaced the fluorophore/donor transfers the energy from its excited state to the quencher/acceptor, rather than fluoresce. This energy transfer is highly distance dependent, making FRET pairs highly accurate at judging small spatial distances; they are often used for DNA tiles, and DNA based computation systems[8,19,20]. However, the cost of fluorophore/quencher conjugation to DNA is high and they are best suited to probing a relatively small region of the nanostructure.

An alternative for probe molecules is adding an intercalating dye, whose quantum yield changes when bound to ssDNA/dsDNA to solution[13,14,16,21]. Intercalating dyes are relatively inexpensive and are convenient to add to existing systems[22–24]. Additionally, entire sets of dyes called High Resolution Melt, or HRM, dyes are designed to detect small differences in DNA[16,25]. The high signal created by dye intercalation allows small amounts of dyes to be used, minimizing their distorting effect on the DNA structures.

Figure 3. A & B shows typical transition data for UV-VIS and fluorescence spectroscopy using intercalating dyes. Figure 3.4 shows SYBR Green I and Rhodamine-X, or

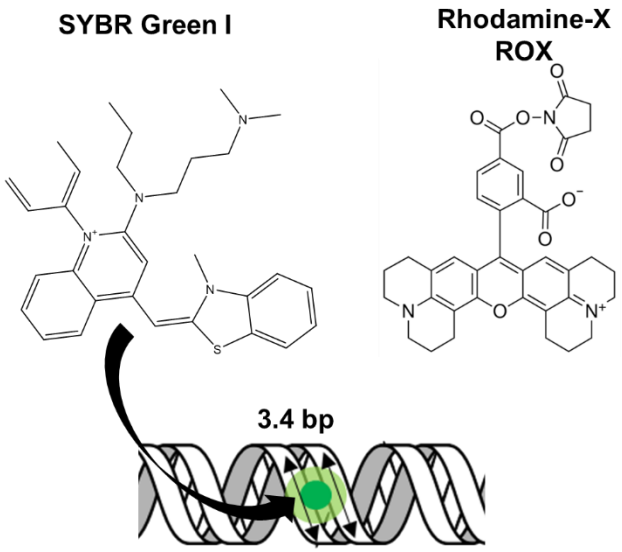


Figure 3.4: Chemical structures of SYBR Green I and ROX fluorescent dyes

ROX, which are typically used intercalating and reference dyes, respectively, for fluorescence systems. SYBR Green binds strongly to dsDNA and takes up approximately 3.4 bases along the dsDNA backbone[16]. While only a first generation High Resolution Melt dye, SYBR Green has a change in fluorescence on binding of >1,000x for dsDNA. SYBR Green I preferentially binds dsDNA to ssDNA, and is primarily enthalpy driven, but has a positive entropy of binding. The positive entropy rises from the release of water or salt molecules during dye binding. This positive entropy ensures that the SYBR Green/dsDNA complex does not destabilize with increasing temperature over the measurable temperature range in water.

Another notable technique is the performance of fluorescence spectroscopy in a Real Time-Polymerase Chain Reaction (RT-PCR) machine. The data obtained from RT-PCR equipment is functionally identical to that from a fluorescence spectrophotometer. However, RT-PCR equipment and sample preparation have significant differences from fluorescence spectrophotometers. Fluorescence spectrophotometers can scan excitation/absorption and emission across wide spectral ranges. This is in enabled by the 90° angle between excitation and emission which prevents the unabsorbed light from being read as emission[26].

Unfortunately, this geometry precludes gathering data from a large number of samples at once. In a RT-PCR machine, a 96 or 384 well plate is laid over a UV excitation source, the emitted fluorescence and UV excitation light both then travel up into the same optics. This allows for massively parallel data gathering, and small sample volumes, but requires optics which redirect emission wavelengths. In an RT-PCR, only the emission wavelengths for which one has installed, and calibrated, a filter may be used. This limits the available reporter molecules and incurs a cost penalty on implementing new dyes. Additionally, RT-PCR equipment often necessitates reference dyes such as Rhodamine-X, or ROX shown in Figure 3.4. Reference dyes are designed to maintain constant fluorescence as a function of temperature and concentration. As RT-PCR samples are so small, reference dyes allow normalization of small changes in concentration due to evaporation or bubbles. Figure 3.3 shows the raw fluorescence of a sample containing no DNA, while Figure 3.3 B shows the reference dye normalized fluorescence of the components in a DNA origami system.

The massively parallel data gathering, low sample volume, and high sensitivity make RT-PCR equipment ideal for testing the thermal properties of DNA nanostructure systems.

In this section, we have discussed common tools to probe the energetics of reversible phase transitions which are relevant to DNA nanostructure systems. We have additionally established the benefit of using RT-PCR equipment to perform anneal/melt experiments. In section 3.4, we will discuss our choice of experimental conditions and compare it to others in the literature.

Section 3.4: Experimental Methods

The fluorescence melt/anneal samples described below were gathered using the Quantstudio 6 system, thanks to the Lommel lab (in NCSU Department of Plant Pathology). Fluorescence samples were, unless otherwise specified, prepared in sets of 3 replicates each

of 20 uL of volume, of 5-15 nM dsDNA, with a ratio of SYBR Green/base pair dsDNA of 1/900 as per Dietz[12], ROX was added in 1:50 ratio by volume, and samples were buffered in pH 5.5 sodium cacodylate buffer with 12.5 mM Mg^{2+}. Sodium cacodylate buffer was chosen as its pH does not vary with temperature as Tris based buffers do, and changes in pH could affect fluorescence intensity and melting/annealing of DNA. Melt and Anneals were performed at an overall rate of 0.21 °C/min, or 5.5 Hr for a single melt or anneal between 25°C and 95°C. As this rate is lower than the software minimum for melt/annealing, the autodelta function with a step/equilibrate cycle of 0.07°C per step with a 20 second equilibration was used, leading to a rate of 0.21°C/min. This approach is similar to the Dietz protocol except in that it was performed at a rate closer to the typical 2-4 hr anneal. This rate of annealing provides more data points for each melt/anneal than standard RT-PCR melt protocols; creating acceptable derivative plots requires Savitzky-Golay smoothing with a relatively large (50-100 point) window. As the equipment software was unable to baseline correct this data we developed and used the python script reported in Appendix 1.

Of the three groups to perform fluorescence spectroscopy on DNA nanostructure systems, all explicitly used RT-PCR to gather fluorescence measurements. As discussed

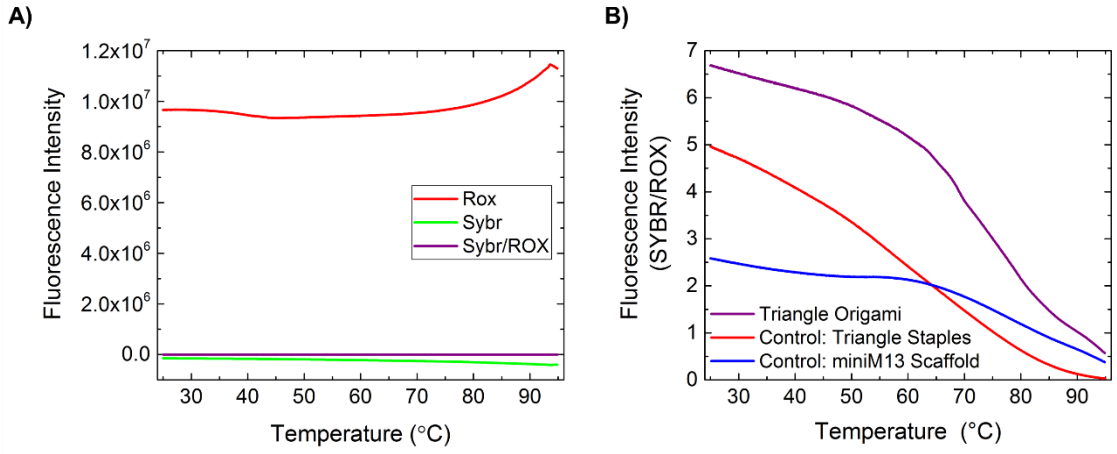


Figure 3.5: Raw fluorescence data, A- dye controls, B-DNA origami components and complete system

above, all of these studies either examined dθ/dT or raw fluorescence rather than θ, with Ouldridge/Turberfield citing the difficulty of baseline correction as the cause[4]. To discuss the baseline corrections necessary to determine θ, it is worth discussing typical RT-PCR samples of biological DNA for which finding θ is well established. In these systems, melting is often a final quality control step and not the primary feature of interest in and of itself[26]. Additionally, melts/anneals for PCR samples often are performed at rates as quick as several °C/min. As DNA nanostructures consist of hundreds of hybridization/denaturing events, a slow temperature ramp is desirable both for maintaining equilibrium and for increased data acquisition points. Dietz, Ouldridge/Turberfield, and Estevez-Torres groups all performed fluorescence based melting/annealing at slow rates, 1°C per 3 hours, 1°C per minute, and 1° per 10-30 min, respectively[4,5,12]. The raw curves in Figure 3.3 B, are nearly identical to the raw curves published in the Dietz work. The lower baseline in Figure 3.3 B appears truncated as the sample concentration, 60 nM, was high enough for very rapid annealing.

Both fluorescence and UV-VIS spectroscopy produce the same figure of merit, θ as a function of temperature, and their baseline correction is nearly identical[13,15]. However, the differences between their mechanisms necessitate deeper discussion. In UV-VIS the changes in absorption of ssDNA and dsDNA can be described by Eq. 3.1 where ε is the extinction coefficient[15]. The upper and lower baselines are defined where the sample is fully hybridized or denatured, allowing baseline correction to solve for θ in between. The baselines are identified as regions outside of the ssDNA/dsDNA transition and are removed by Eq. 3.2 or Eq.3.3 for UV-Vis and fluorescence respectively[14]. However, the baselines for fluorescence data are caused by the change in quantum yield of the dye as a function of temperature. This change in quantum yield is attributed to intermolecular collisions that drop dye molecules from

the excited state without emitting photons. Such collisions are driven by diffusion, making quantum yield proportional to KT.

3.1) $A(T) = \varepsilon_{ss}(T)\theta_{ss}(T) + \varepsilon_{ds}(T)\theta_{ds}(T)$

3.2) $\theta_{ssDNA}(T) = \frac{A(T)-A_L(T)}{A_U(T)-A_L(T)}$

3.3) $\theta_{dsDNA}(T) = \frac{F(T)-F_L(T)}{F_U(T)-F_L(T)}$

Eq. 3.1, 3.2 and 3.3 describe the absorption, fraction ssDNA, and fraction dsDNA as a function of temperature respectively, all for UV-VIS and fluorescence experiments. F_U,A_U,F_L and A_L are all functions fitted over their respective regions.

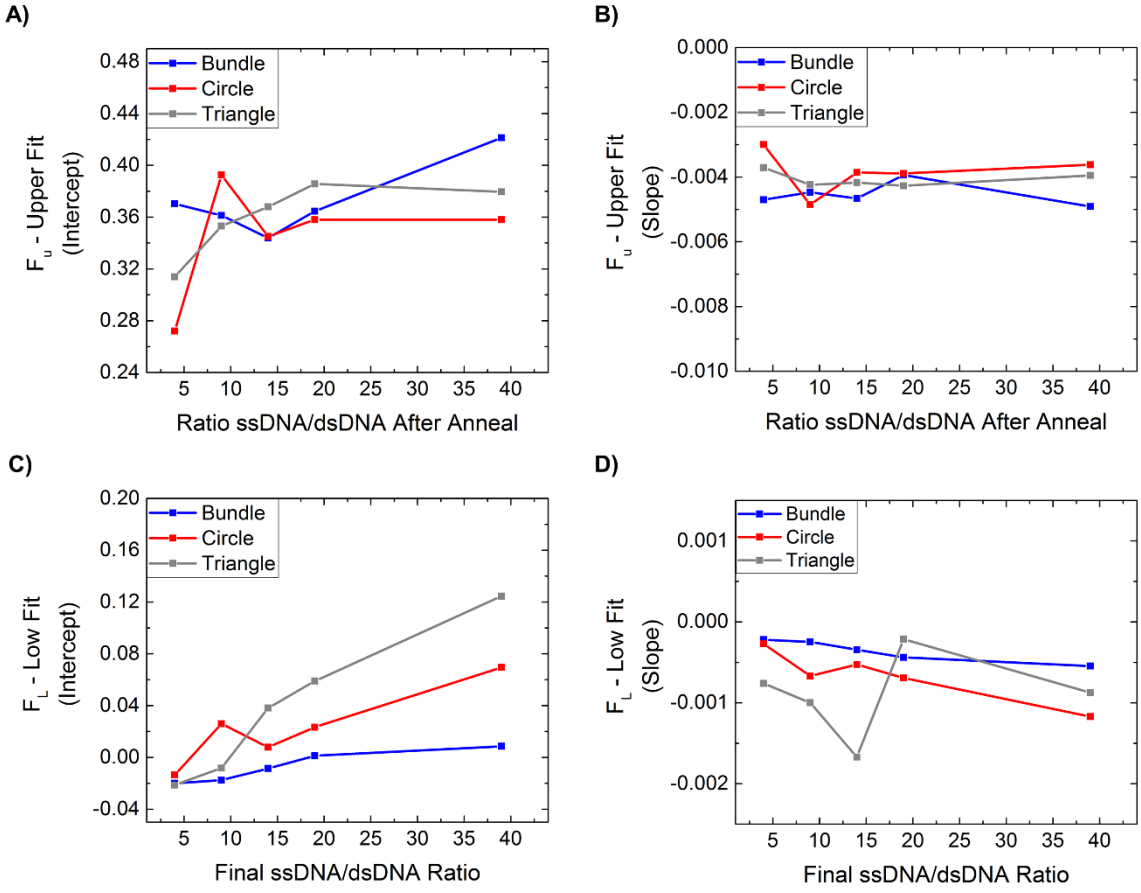


Figure 3.6: F_U and F_L fits for scaffolded DNA nanostructures as a function of excess ssDNA

Another distinction for intercalating fluorescence dyes, particularly SYBR Green I, is their interactions with ssDNA. It should not be assumed that intercalating fluorescent dyes bind equally to ssDNA and dsDNA[16]. SYBR Green in particular will preferentially bind dsDNA over ssDNA. Evidence of this is visible in reference dye normalized fluorescence data, a qualitative examination of curves in the presence of ssDNA shows that the total fluorescence at low temperature cannot reasonably be the sum of ssDNA and dsDNA contributions. Additionally, ssDNA fluorescence curves are often discussed in RT-PCR trouble shooting guides, but not in baseline correction[26]. A basic test of this was to examine the upper and lower fluorescence fits, F_U and F_L while increasing excess ssDNA. Figure 3.6 shows the changes in F_L and F_U as a function of excess ssDNA for a set of three miniM13 origami structures annealed with staple concentrations of 5x, 10x, 15x, 20x, and 40x. It shows a clear

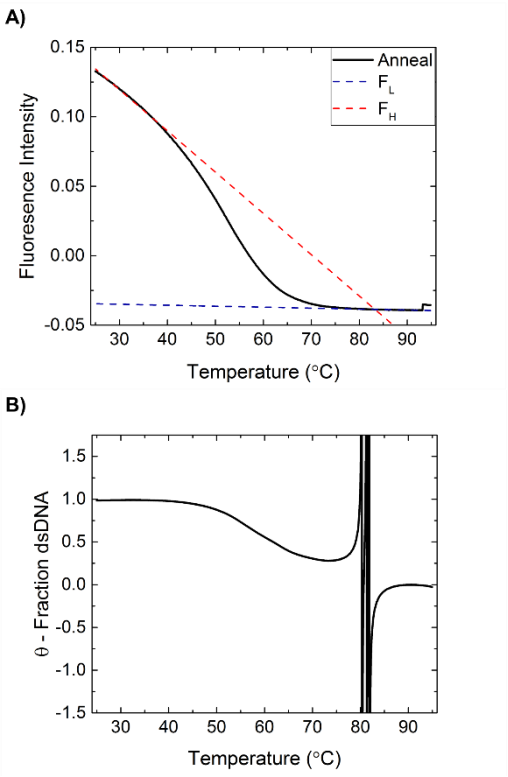


Figure 3.7: Reference dye normalized fluorescence and linear baseline corrected θ for annealing 5 nM bundle at 5x staple excess

dependence on excess ssDNA for F_L, particularly in its intercept, but no dependence for F_U. This distinction plays a vital role when testing samples across a variety of concentrations and excess ssDNA. Increasing the excess ssDNA to drive nanostructure formation minimally affects the signal to noise ratio in fluorescence spectroscopy compared to UV-VIS.

In this section, we have discussed the experimental procedure we used to generate fluorescence spectroscopy annealing/melting curves, how those relate to ones used in the literature, and why the signal to noise ratio is not reduced by excess ssDNA in a SYBR green system. We have also briefly touched on the theory of baseline corrections for intercalating dye systems, such as SYBR green.

Section 3.5: Baseline Correction of Fluorescence Data

One caveat to the differences between UV-VIS and fluorescence data for melting is their ability to form vertical asymptotes. For both Eq. 3.2 and 3.3, the corrected θ will reach an asymptote if $F_L = F_H$ or $A_L = A_H$. Figure 3.7 illustrates why this could be more of a concern for fluorescence data; the difference in slope between F_L and F_H makes their intersection close to the region of interest. Figure 3.7 shows an example annealing curve, 5 nM bundle origami at 5x staples, before and after baseline correction. The location of the vertical asymptote matches intersection of the fit lines at 81.3°C. The relative location of F_L=F_H explains why baseline correction is a delicate affair for low concentration samples; curvature in the F baseline, poor selection of fitting windows, or other factors can mask the corrected data with difficult to remove vertical asymptotes

Given the similarity between the protocols used here and those used in previously for DNA nanostructures, described in section 3.2, it is unsurprising that baseline correction was non-trivial. On visual inspection, the melt and anneal curves contain the appropriate regions and transitions to perform baseline correction. However, only a moderate proportion of the

anneal curves, and a minority of the melt curves, ere usable after linear baseline correction. The vertical asymptotes, typically in the transition or in the F_L region, or of excessive curvature in either the F_L or F_U regions, made further analysis impossible for these systems. The higher probability of a vertical asymptote in melt data will be addressed separately in section 3.6.

As the nanostructure fluorescence data, shown in both Figure 3.7 B and Figure 3.7 A, visibly contains the appropriate regions and transition, it is reasonable to assume that the problems occur within data processing. To address this, alternative baseline corrections were investigated. Ultimately, a thus far unimplemented, method developed by the Wittwer lab[23]. This method uses a temperature dependent expansion of beer's law for addressing

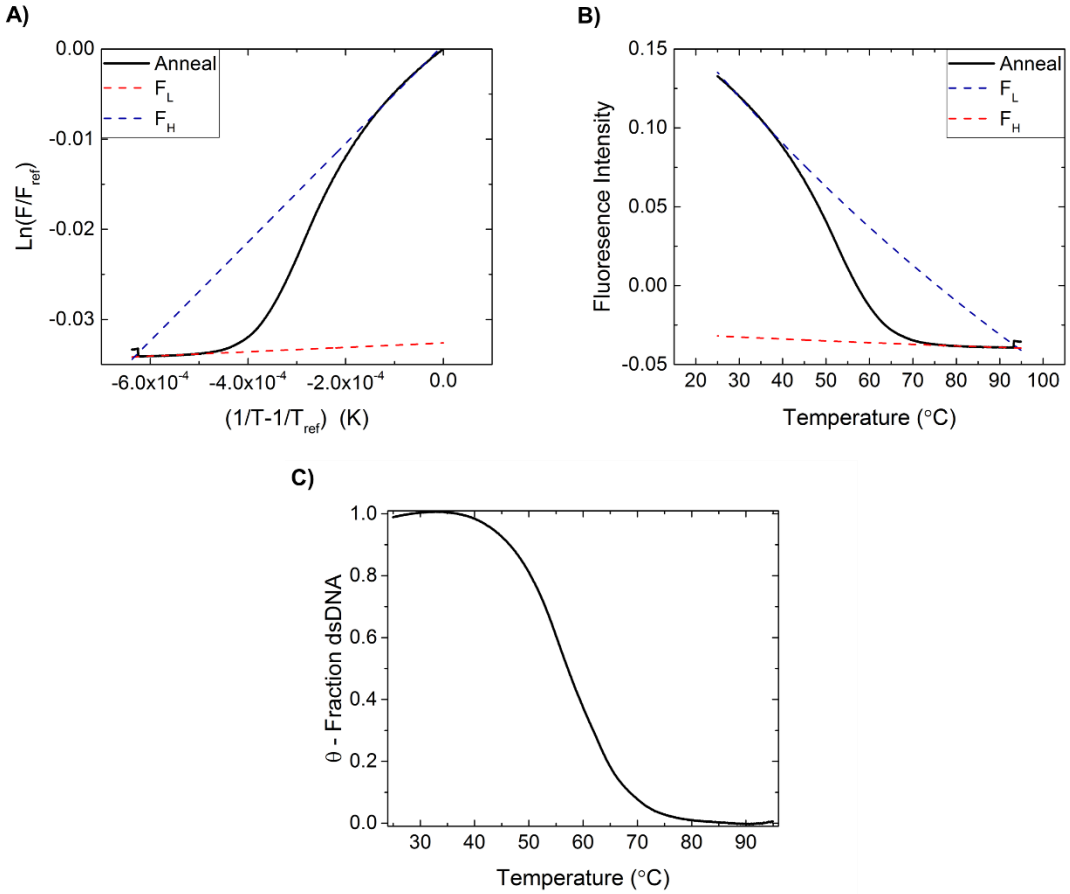


Figure 3.8: Baseline correction using Wittwer function, A-data transformation for linear fitting, B-fit functions in typical units, C-baseline corrected data

fluorescence[27]; arguing that a baseline of the form in Eq. 3.4 is more in keeping with the mechanisms for fluorescence spectroscopy.

3.4) $$\ln\left({}^{F}/_{F_{ref}}\right) = C\left(\frac{1}{T} - \frac{1}{T_{Ref}}\right)$$

Figure 3.8 A shows transformation of the data into coordinates appropriate to fit Eq. 3.4, with the subsequent fit lines. Return transformation shows the intersection point where $F_L = F_H$ to be at a higher temperature for this Beer's law baseline than for the standard linear baseline. Figure 3.8 C shows the final baseline corrected data for the 5 nM bundle nanostructures, clearly without the asymptotes that result from linear baseline correction.

Section 3.6: Photobleaching

Despite the significant reduction in vertical asymptotes by using the Wittwer baseline, the propensity for melt data to form vertical asymptotes is concerning. These asymptotes are relatively rare for biological samples it is likely that they indicate an additional artifact in the fluorescence data. To illustrate the differences between melting and annealing data, the weave tile was chosen. The weave tile is a simple structure of two ssDNA strands combing to form two helixes of dsDNA 16 base pairs long, which feed into each other at the ends. The weave tile nanostructure, is shown in Fig 3.4 A.

Figure 3.9 shows fluorescence melt and anneal curves, in the Mg rich buffers described in section 3.4. Visual inspection of the melt and anneal curves in Figure 3.9 B yields several important insights. The slope in the upper fit of the melt is steeper, reducing the temperature at which it intersects the lower fit and creates a vertical asymptote. This difference in slope for reference dye normalized fluorescence is consistent across all nanostructures.

The region below 50°C for both the melt and anneal curves should contain identical amounts of dsDNA. Even in lower ionic strength wash buffer, UV-VIS indicates a Tm of

57.5°C, Appendix 2. As such, the difference in fluorescence for the melt and anneal curves in the higher ionic strength TAE Mg buffer is likely an experimental artifact.

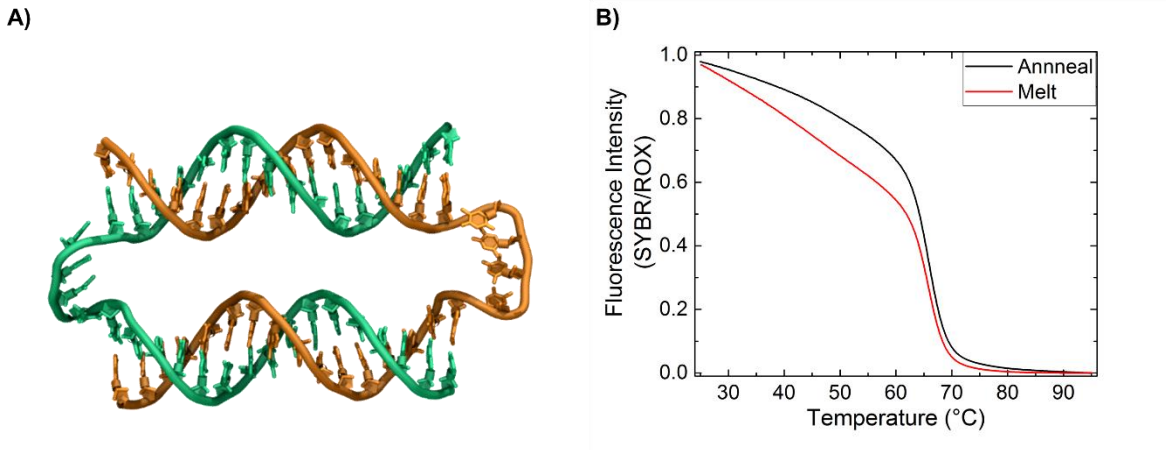


Figure 3.9: Weave tile visualization and anneal/melt curves

Another interesting feature of Figure 3.9 B is the slight convex curvature in the anneal curve compared to the linear melt. The F_U slope of both curves should be identical and directly related to quantum yield of the SYBR Green when bound to DNA, barring some chemical change to the dye or buffer.

3.5) $$I_t = I_0 e^{-\left(\frac{t}{\tau}\right)^{\left(1/h\right)}}$$

A likely candidate for these artifacts is photobleaching of the SYBR Green dye. Photobleaching of dyes is an exponential process as a function of exposure time[28], as described in Eq. 3.5, where t is exposure time, τ is the characteristic time scale of the decay process, and h is a homogeneity factor that relates to the ability of the buffer to absorbe the UV excitation light. In this context t should be directly proportional to both cumulative data acquisition points and to temperature. In the context of fluorescence microscopy with the same SYBR Green I dye, the homogeneity factor represents the likelyhood that UV excitation will be absorbed by the cellular material around the dye/DNA complex. In this context the

homogeneity factor should be relatively constant due to the consistency of buffer conditions across samples.

Photobleaching is often checked for, and not observed, in typical RT-PCR fluorescence systems[16,22,29]. However, as described in section 3.4, the protocols for DNA nanostructures necessitate a higher number of UV exposures/data acquisitions which could make photobleaching more likely. The cumulative UV exposure time as a function of temperature will progress in opposite directions for the melt and anneal ramps; this could account for the differences in slope/curvature between melt and anneal curves. That the anneal occurs first, and the melt has a lower intensity, further indicates photobleaching.

Two tests were performed to identify photbleaching of intercalated SYBR Green, shown in Figure 3.10. In the first, weave tiles were annealed, melted, and annealed over four cycles, over approximately 38 hours. While the sample was exposed to UV excitation during

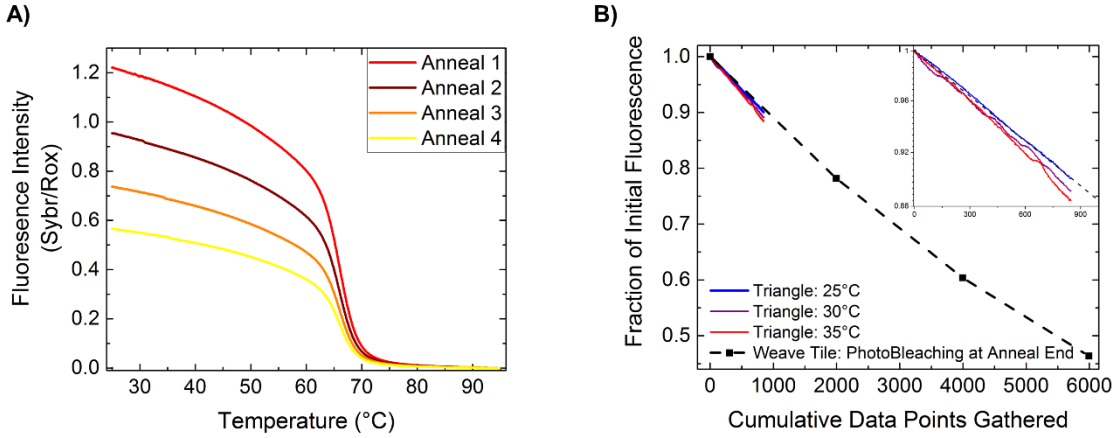


Figure 3.10: Photobleaching shown in A- consecutive anneals of weave tile, B- continuous isothermal photobleaching with representative points from A

melts, the emission was not recorded for technical reasons. Figure 3.10 A, shows the resulting anneal curves, with a clear reduction in fluorescence over time.The second test was to hold three replicates of various DNA structures at constant temperature with the same data sampling rate as a typical melt or anneal ramp. This was repeated on the same RT-PCR plate

of samples for 25°C, 30°C, and 35°C. The fluoresence was then plotted as a fraction of the maximum initial value. One such system, the triangle miniM13 origami, is compared to the photobleaching of the weave tile in the first experiment, shown in Figure 3.10 B. There is reasonable agreement between the two curves, while the inset shows a moderate temperature dependence for the isothermal curves. Both sets of data show the slight curvature one would associates with plotting an exponential decay on a scale similar to its half life.

Removal of photobleaching effects first requires the use of Eq. 3.5 to predict photobleaching for any given point in a melt or anneal ramp. Removal of photobleaching also requires the use of controls to determine the accuracy of baseline corrected values for ssDNA/dsDNA content. This poses several complications and will be discussed more thoroughly in future work. A truly exhaustive exploration of photobleaching rates would require examination of temperature, pH, buffer molecule, and ssDNA concentration. The following set of experiments represent a preliminary examination of this parameter space.

Initial exploration of parameter space was performed via 5.5 hr isothermal photobleaching runs at 25°C, 30°C, and 35°C. Five DNA nanostructures, two buffer types, two buffer pH values, two nanostructure concentrations and two scaffold sizes were chosen

Table 3.2: Isothermal photobleaching sample set. Tall Rectangle is abbreviated TR.

Sample	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***
Concentration (nM)	5	5	**2.5**	2.5	15	**7.5**	15	15	15	15	15
pH	7.4	7.4	7.4	**5.5**	7.4	7.4	7.4	**5.5**	5.5	5.5	5.5
Buffering Agent	NaCac	**Tris**	NaCac	NaCac	NaCac	NaCac	**Tris**	NaCac	NaCac	NaCac	NaCa
Structure	TR	TR	TR	TR	Circle	Circle	Circle	Circle	Bundle	Triangle	F16

for the 11 samples described in Table 3.6. Each sample had 3 replicates in the photobleaching runs, resulting in a total of 99 photobleaching curves. These samples were chosen specifically to answer whether any one factor had a significant enough role in photobleaching to require inclusion in the correction factor. Nanostructure size and shape account for a large proportion

of the variation in the 11 samples, as any successful correction must be effective for arbitrary nanostructures.

The photobleaching experiments were processed, and the last 800 data points of each curve was fit to Eq. 3.5. The first 200 data points were masked to avoid fluctuations associated with the temperature increments. Fitting data, and goodness of fit is shown in Appendix 3.

3.6) $$I_{t_{Melt}} = I_0 e^{-\left(\frac{T*C}{\tau}\right)^{\left(1/h\right)}}$$

3.7) $$I_{t\_Anneal} = I_0 e^{-\left(\frac{(T_{max}-T)*C}{\tau}\right)^{\left(1/h\right)}}$$

3.8) $$C = \frac{Data\ Points\ per\ Curve}{T_{max}-T_{min}}$$

Figure 3.11 A shows a plot of every fitted τ/h parameter pair. For the purposes of a general photobleaching correction, the most important test is whether a single τ/h pair in Eq.

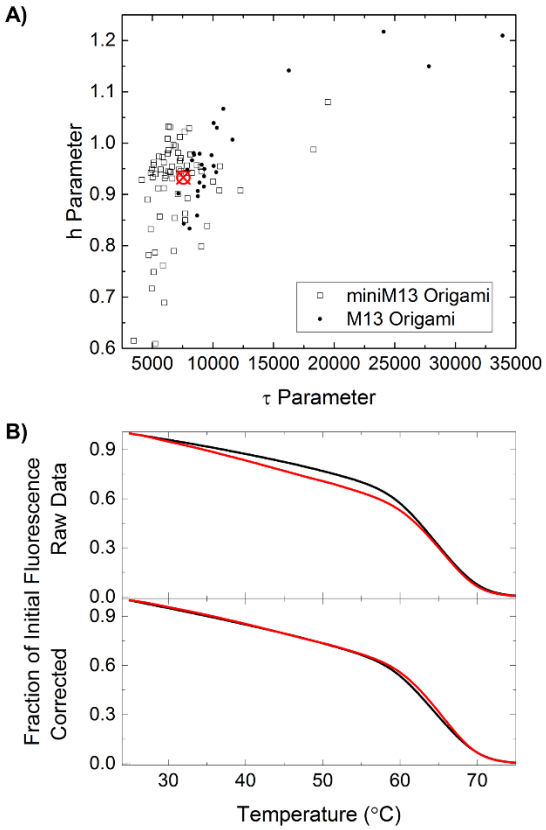


Figure 3.11: A-all fitted photobleaching parameters, B- raw data correction for the weave tile based on average parameter (Red crosshair in A)

3.6 & 3.7 can correct the data well enough without accounting for variance in τ/h as a function of other factors. Reasonable criterion for correcting the data well enough is overlap of upper fit region and final θ data which matches hysteresis and T_m for samples without photobleaching.

Figure 3.11 B, which shows the raw weave tile data before and after correction with Eq. 3.6 & 3.7 with a static τ/h pair, indicated by the red crosshair in Figure 3.11 A. This τ/h pair is the averaged τ/h for all the fitted photobleaching curves.

The weave tile data, gathered in wash buffer for previous work, discussed above is not sufficient to provide a check for the hysteresis/T_m criteron. Evaluation of those critereon will require performing melt/anneal data with UV, and with fluorescence at data aquisiton rates that will not induce photobleaching.

Figure 3.12 and Figure 3.13 address the preliminary probe of the contributing factors to photobleaching for temperature and pH respectively. From Figure 3.13 A, C and D it is evident that there is a temperature dependence of the photodecay characteristic timescale on temperature. A literature model for the temperature dependence of similar processes should be found, and fitted to this data to compensate for changes in temperature during anneals.

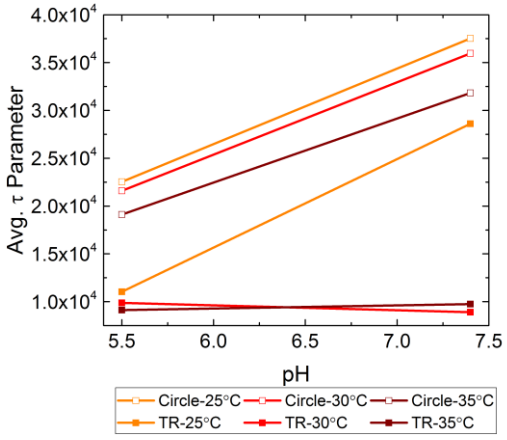


Figure 3.12: τ parameter fits as a function of pH

Given that photobleaching and fluorescence change based on the number of intermolecular collisions, viscosity may also be a meaningful parameter to examine.

These factors altogether, along with excess ssDNA in the buffer solution must be addressed in order to perform accurate analysis of hysteresis and melting/annealing transitions.

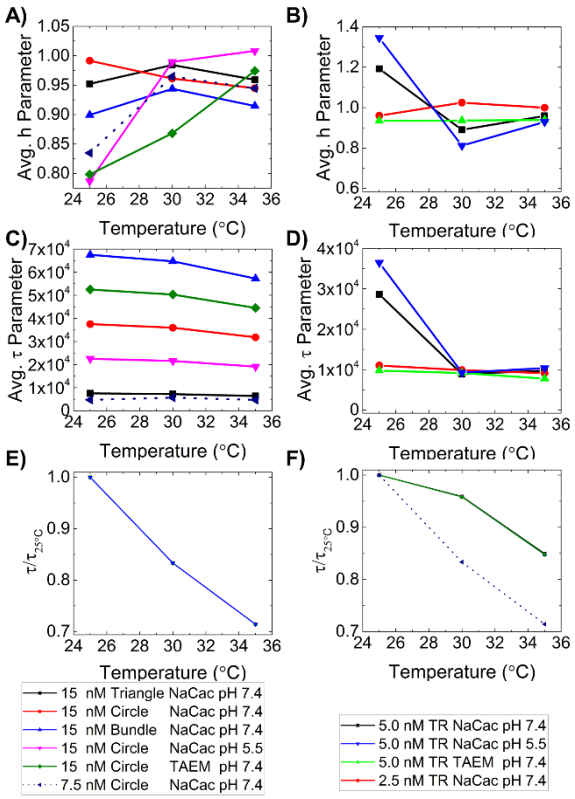


Figure 3.13: h and τ parameter fits (A-B, and C-F respectively) as a function of temperature for miniM13 (A,C,E) and m13 (B,D,F) structures. τ normalized by the initial value for E &F.

Section 3.7: Application to DNA Nanostructures

. As the corrections for photobleaching still require improvement, and the interaction of photobleaching corrections with baseline correction are unexamined, photbleaching corrections were ommited from the following analysis of DNA origami systems. The error in

hysteresis measurements should be reasonably constant between samples, particularly for those with a high signal/noise ratio like the higher concentration m13 DNA nanostructures.

The fitting windows used for F_L and F_H were manually selected, rather than implementing the second derivative method described by owcazarzy[14]. Implementation of the second derivative fitting window selection was left for future work; We were concerned with programming overhead for performing the necessary savitzky-golay smoothing, and with the potential complications from the more gradual melt/anneal transition of nanostructures which would make the second derivative more subtle. A set of preselected fitting windows were identified by examining the transitions in the reference dye normalized SYBR Green I fluorescence. The baseline corrected data was evaluated for vertical asymptotes and for curvature in the fitting regions.

The Wittwer baseline correction was largely sucessful for the model M13 origami, theTall Rectangle, or TR. Vertical asymptotes were not observed in TR anneal or melt curves, nor was there undesirable curvature in F_H or F_L regions. Further, the hysteresis was both qualifiable visually and quantifiable by area, as illustrated by Figure 3.14. The hysteresis area between the 1 nM and 12 nM systems clearly decreases, with an increase in the temperature

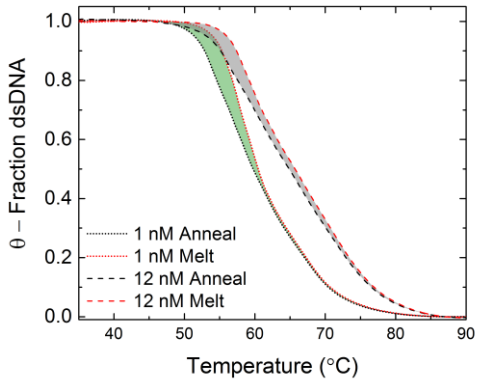


Figure 3.14: Hysteresis area for Tall Rectangle nanostructures at 1 and 12 nM concentration

of transition. The breadth of the transition also increases with increasing scaffold concentration. This illustrates the importance of appropriate baseline correction; we are able to probe the coopertivity through the hysteresis area and breadth of transition as a function of processing conditions.

As described in chapter 1.3 the contribution of coopertivity increases the free energy of transition for annealing, but decreases it for melting. This drives the melt and anneal curves to higher and lower temperature respectively. The driving force contributed by scaffold concentration and staple excess should push both curves to a higher transition temperature. As other contributions to driivng forces increase, such as ionic strength or DNA concentration, the relative contribution of the coopertivity should be less, resulting in a lower hysteresis area.

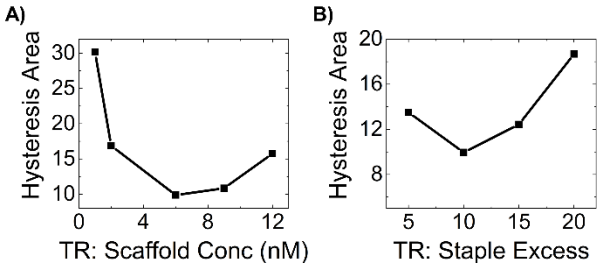


Figure 3.15: Hysteresis area as a function of scaffold concentration and staple excess

To more thoroughly examine the hysteresis area as a function of driving force, the TR was annealed at scaffold and staple concentrations of a typical anneal, 10 nM scaffold with 10x excess staples and staple excess concentration. The resulting hysteresis area is plotted in Figure 3.15. Both scaffold concentration and staple excess create similar trends in the measured hysteresis area; this is reasonable as both drive the annealing process through reagent concentration. The curvature in these plots is particularly interesting; it implies that while the relative contribution of coopertivity is decreased with increasing concentration, there is also dependence of the coopertivity on concentration. From the limited data set available,

shown in Appendix 3, we observe that this increase in area coincides with a slight increase in melt curve transition while the anneal curve remains constant.

While the baseline correction, in particular photobleaching correction, must be improved and more data must be gathered, it is clear that melt/anneal curves and hysteresis

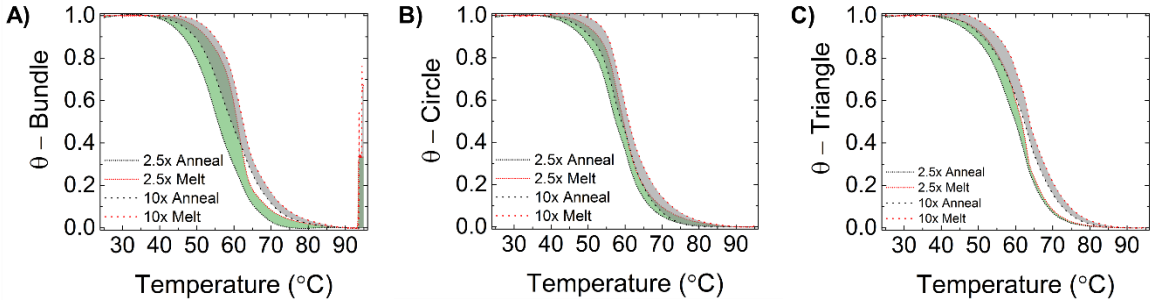


Figure 3.16: Hysteresis plots for miniM13 structures A-bundle, B-Circle, C-Triangle at 5 nM and staple excesses of 2.5x and 10x

area have the potential to reveal much about DNA nanostructure assembly.

Beyond probing the driving forces involved in nanostructure self assembly, θ as a functio nof temperature can help in understanding the differences between structures and potential design choices within structures. Figure 3.16 A, B & C, shows the θ melt/anneal curves and hysteresis for 15 nM miniM13 origami with 2.5x and 10x staple excesses respectively. The bundle origami, Figure 3.16 A, has a vertical asymptote at ~93°C which does not visibly shift the melt/anneal curves. Overall, this data set provides useful information for comparing these nanostructures. The bundle, as expected for a 3D origami, has a much lower annealing temperature, and a higher hysteresis area compared to either the circle or the triangle. The similarity between the planar circle and triangle origami, and the overall trend

of decreasing hysteresis for higher transition temperatures are other features which could not be easily be determined from dθ/dT.

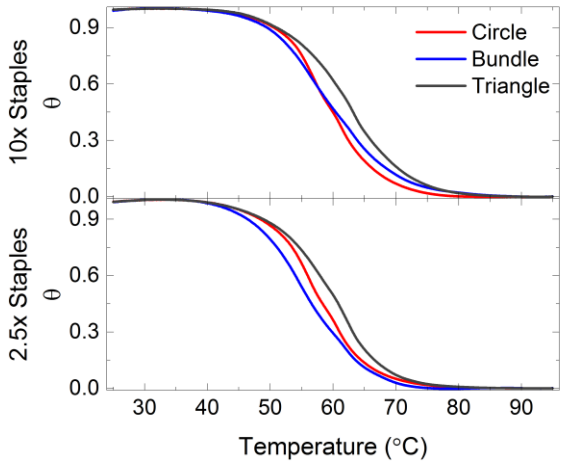


Figure 3.137: Annealing curves for Circle, Bundle, and Triangle nanostructures

For the purposes of processing, comparison between annealing curves is also interesting. Figure 3.117 shows the annealing curves of the bundle, circle, and triangle for 10x and 2.5x excess staple concentrations. Of particular interest is the differences in how much each curve shifts as a function of excess staples; while the triangle is clearly the fastest to assembly in both cases, if the circle and bundle were to compete at 2.5x and 10x excess staple concentration, the results would be quite different.

Vertical asymptotes were more common for miniM13 origami, in melt curves, and for the bundle structure. This makes accurate analysis of hysteresis area much less reliable. Once a complete photobleaching procedure is defined and implented within baseline correction we will be able to perform more in depth analysis of the driving forces in these miniM13 nanostructures.

Section 3.8: Conclusions and Future work

In this chapter we have established the importance of, in section 3.1, thermal properties for the future implementation of DNA nanostructures; in section 3.2, the fraction of

dsDNA, or θ, as a function of temperature for DNA nanostructures and how it can be used to probe relavent properties; in section 3.3, the variety of techniques which can probe these properties and the appeal of RT-PCR instrumentation; in section 3.4, we defined our standard experimental conditions in the context of previous work both on DNA nanostructures and in biological systems; in section 3.5, we showed that the use of the Wittwer baseline function significantly improves standard corrections for nanostructure fluorescence data; in section 3.6, we confirmed the presence of photobleaching artifacts and showed preliminary steps for removal of photobleaching artifacts; in section 3.7, we implemented the Wittwer baseline function and demonstrated the utility of θ curves for DNA nanostructures.

Bringing this work to completion will require completion of several sequential experiments. First and foremost, the selection of fitting windows for the baseline correction will be automated as per Owcazarzy[14]. This should improve the consistency of baseline correction, and remove a source of human error. Secondly, the contributing factors to photobleaching will be further examined; In particular, the photobleaching characteristic timescale, τ as a function of temperature and pH and, the homogeneity factor, h as a function of excess ssDNA. These factors will be either fitted to a literature model, or ruled out as contributors to photobleaching. Thirdly, multiple control samples for determining removal of photobleaching both in transition temperature and hysteresis area must be analyzed to confirm correction of the photobleaching artifact.

Finally, two other consideration must be made regarding the extensibility of these methods to other laboratories. The first being that quantification of hysteresis should be automated, and figures of merit for the breadth of melt/anneal transition should be adopted. The last potential point of improvement is that of second generation high resolution melt dyes.

While SYBR Green has been the most common dye in the literature, adoption of a newer, more accurate system such as SYTO 9 is worthy of consideration.

By developing our experimental thermodynamics toolkit for DNA nanostructures, our ability to understand and improve their formation will benefit. Ultimately, such a foundation will hasten the growth of DNA nanotechnology as an effective and low cost nanofabrication technique.

References

1. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. *Nucleic Acids: Structures Properties, and Functions*. (University Science Books, 2000).

2. SantaLucia, J. & Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **33,** 415–440 (2004).

3. Lu, M., Guo, Q., Marky, L. a, Seeman, N. C. & Kallenbach, N. R. Thermodynamics of DNA branching. *J. Mol. Biol.* **223,** 781–789 (1992).

4. Dannenberg, F., Dunn, K. E., Bath, J., Turberfield, A. J. & Ouldridge, T. E. Modelling DNA Origami Self-Assembly at the Domain Level. 1–22

5. Estevez-torres, A. Observing and Controlling the Folding Pathway of DNA Origami at the Nanoscale. (2016). doi:10.1021/acsnano.5b05972

6. Ke, Y., Bellot, G., Voigt, N. V., Fradkov, E. & Shih, W. M. Two design strategies for enhancement of multilayer–DNA-origami folding: underwinding for specific intercalator rescue and staple-break positioning. *Chem. Sci.* **3,** 2587 (2012).

7. Martin, T. G. & Dietz, H. Magnesium-free self-assembly of multi-layer DNA objects. *Nat. Commun.* **3,** 1103 (2012).

8. Wei, X., Nangreave, J., Jiang, S., Yan, H. & Liu, Y. Mapping the thermal behavior of DNA origami nanostructures. *J. Am. Chem. Soc.* **135,** 6165–6176 (2013).

9. Arbona, J.-M., Aimé, J.-P. & Elezgaray, J. Cooperativity in the annealing of DNA origamis. *J. Chem. Phys.* **138,** 015105 (2013).

10. Arbona, J. M., Elezgaray, J. & Aimé, J. P. Modelling the folding of DNA origami. *EPL (Europhysics Lett.* **100,** 28006 (2012).

11. Dunn, K. E. *et al.* Guiding the folding pathway of DNA origami. *Nature* (2015). doi:10.1038/nature14860

12. Sobczak, J.-P. J., Martin, T. G., Gerling, T. & Dietz, H. Rapid folding of DNA into nanoscale shapes at constant temperature. *Science* **338,** 1458–61 (2012).

13. You, Y., Tataurov, A. V & Owczarzy, R. Measuring Thermodynamic Details of DNA Hybridization Using Fluorescence. *Biopolymers* **95,** 472–486 (2011).

14. Owczarzy, R. Melting temperatures of nucleic acids : Discrepancies in analysis. **117,** 207–215 (2005).

15. Marky, L. A. & Breslauer, K. J. Calculating Thermodynamic Data for Transitions of any Molecularity from Equilibrium Melting Curves. **26,** 1601–1620 (1987).

16. Paper, O. SYBR Green I : Fluorescence Properties and Interaction with DNA. 1189–1199 (2012). doi:10.1007/s10895-012-1059-8

17. Harris, D. C. Applications of Spectrophotometry. *Quant. Chem. Anal.* **(8th ed.) ,** 419–444 (2010).

18. Selvin, P. R. The renaissance of fluorescence resonance energy transfer. *Nat. Struct. Biol.* **7,** 730–734 (2000).

19. Nangreave, J., Yan, H. & Liu, Y. Studies of thermal stability of multivalent DNA hybridization in a nanostructured system. *Biophys. J.* **97,** 563–571 (2009).

20. Hakker, L., Marchi, A. N., Harris, K. A., Labean, T. H. & Agris, P. F. Structural and thermodynamic analysis of modified nucleosides in self-assembled DNA. *J. Biomol. Struct. Dyn.* (2013). doi:10.1080/07391102.2012.763184

21. Wei, X., Nangreave, J. & Liu, Y. Uncovering the self-assembly of DNA nanostructures by thermodynamics and kinetics. *Acc. Chem. Res.* **47,** 1861–1870 (2014).

22. Sanford, L. N. & Wittwer, C. T. Monitoring temperature with fluorescence during real-time PCR and melting analysis. *Anal. Biochem.* **434,** 26–33 (2013).

23. Sanford, L. N., Kent, J. O. & Wittwer, C. T. Quantum Method for Fluorescence Background Removal in DNA Melting Analysis. *Anal. Chem.* 9907–9915 (2013). doi:10.1021/ac4024928

24. Reed, G. H., Kent, J. O. & Wittwer, C. T. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Pharmacogenomics* **8,** 597–608 (2007).

25. Monis, P. T., Giglio, S. & Saint, C. P. Comparison of SYTO9 and SYBR Green I for real-time polymerase chain reaction and investigation of the effect of dye concentration on amplification and DNA melting curve analysis. *Anal. Biochem.* **340,** 24–34 (2005).

26. Bustin, S. A. *A-Z of Quantitative PCR*. (International University Line, 2004).

27. Lemoine, F., Antoine, Y., Wolff, M. & Lebouche, M. Simultaneous temperature and 2D velocity measurements in a turbulent heated jet using combined laser-induced fluorescence and LDA. *Exp. Fluids* **26,** 315–323 (1999).

28. Okabe, K. *et al.* Intracellular temperature mapping with a fluorescent polymeric thermometer and fluorescence lifetime imaging microscopy. *Nat. Commun.* **3,** 705–709 (2012).

29. Zipper, H., Brunner, H., Bernhagen, J. & Vitzthum, F. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Res.* **32,** (2004).

30. TA Instruments: Microcalorimetry Promotional Materials

31. Protocol, B. & Protocol, B. Calorimetry of Nucleic Acids. 1–9 (2000).

32. Gill, P., Moghadam, T. T. & Ranjbar, B. Differential Scanning Calorimetry Techniques : Applications in Biology and Nanoscience. 167–193 (2010).

Chapter 4: Competitive Annealing and Multishape Nanostructures

Section 4.1: Introduction & Motivation

As we discussed in the section 1.4 of the introduction, the sheer number of cooperative events in the self-assembly of a DNA nanostructure can result in emergent behavior. DNA nanostructures consist of tens to hundreds of staple strands, where each staple binds to at least two other sections of DNA. The degree to which these events are cooperative and the degree to which they are independent has not been fully characterized. This results in a hugely complex process or hundreds or thousands of individual hybridization and dehybridization events, each effecting its neighbors[1,2,3]. The ability of these DNA nanostructure systems to form kinetic traps[4], with subsequent yield decreases, can trace its roots back to the significant number of individual events required for formation.

Probing a system from initial ssDNA to a final nanostructure has been a major point of inquiry for DNA nanotechnology[5–10]. As techniques to probe the formation of DNA nanostructures improve, so does our ability to create complex, and potentially revolutionary new tools[1,7,9,11]. In chapter 5, we will discuss those which allow studies of the structure pathway itself. In this chapter we will address techniques with modified conditions which can be imaged at standard temperatures. The nature of these conditions will allow insight into the folding pathways of the structures involved. On such technique, the Turberfield probe of kinetics using conjoined dimer scaffolds, is of particular interest for studying folding pathways.

Figure 4.1, reproduced under fair use, illustrates the dimer scaffold experiment. The standard M13 scaffold molecule was replaced with a dimer consisting of two repeats of the M13 sequence. This gave each staple binding pad multiple choices for hybridization locations. By observing the statistical prevalence of the resulting nanostructures the Turberfield group was able to probe binding preferences of the individual staples. This in turn gives insight into

binding when the staples have only one binding location. The results of this study confirmed the cooperative nature of origami; the dimer structures favoring binding steps which maximized cooperative neighbor-neighbor relations and minimized stepwise reductions in conformational energy. Another description of this tendency is that staples bridged the dimer scaffold at positions which minimized the distance between binding sections along the scaffold backbone.

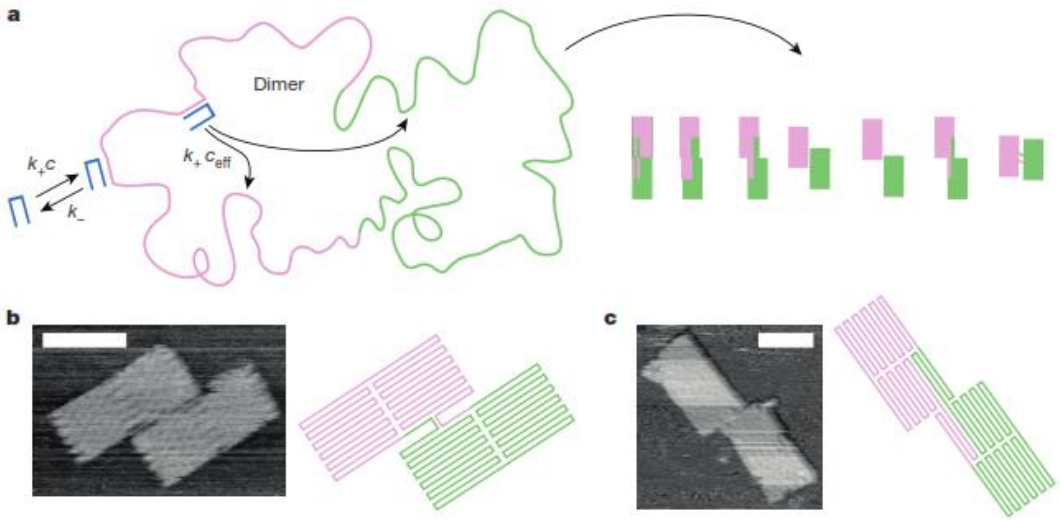


Figure 4.1: Dimer scaffold from Turberfield group

Another thrust of DNA nanotechnology research addresses the degree to which DNA nanostructures fold in a prescribed pathway, and the degree to which their formation can be described as a thermodynamic funnel as do proteins[12,13]. While there is some consensus that the folding process is more path like than protein folding, the probability of multiple discrete dimer origami in the Turberfield work, as well as other studies, indicate that the folding process does not occur purely on one path[1,9]. Processing factors such as concentration or ionic strength could push the structure towards a funnel or a pathway. This is particularly likely if the favorability of all hybridization/annealing events is increased significantly compared to the forces which distinguish events such as conformational entropy and cooperativity. At

particularly high salt or concentration, the relative differences in energy between binding with and without cooperativity could be unable to enforce an ordered pathway.

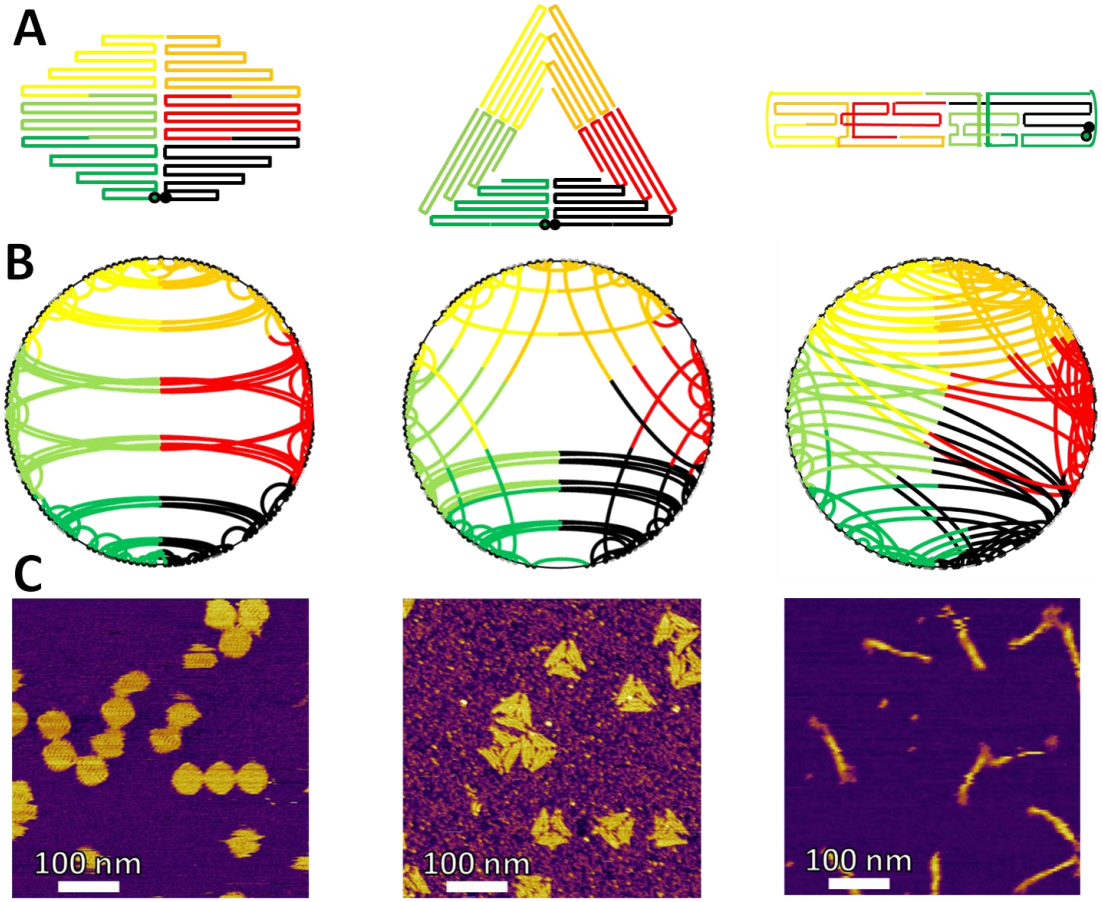


Figure 4.2: Circle, Triangle and Bundle miniM13 DNA nanostructures, A- Routing Map, B- Circle Map, C- AFM data

To better understand the processes driving folding we developed a set of experiments which parallel and build upon the Turberfield work. Rather than creating a new scaffold, we chose to put the staples of differing nanostructure shapes into competition. This setup has the additional benefit of creating new scaffolded nanostructures without additional design or upfront costs. As M13 origami have as many as 220+ staples, and the miniM13 origami have a much more manageable ~64-70 staples, this is a valuable application for the miniM13.

Figure 4.2 shows the three DNA nanostructures that were used to test competitive nanostructure anneals, namely the circle, triangle and 6 helix bundle structures. The 6-helix

bundle being a 3D structure consisting of two 3 helix layers. The triangle structure was designed and ordered after our initial experiments; Its threefold rotational symmetry was chosen to make it easy to distinguish from 6-helix bundle and circle. Figure 4.2 A, depicts the routing paths of the structures, and Figure 4.2 B shows the circle maps of the structures. As described in section 1.4 the outer circle of the map represents the circular ssDNA scaffold, and the arcs inside bridge the positions which a staple will bind together. Larger arcs indicate staples which will reduce the conformational entropy of the scaffold. Figure 4.2 C shows AFM images of the pure nanostructures.

In this section we have addressed both the need and background of techniques which modify standard annealing conditions such that the final state of the system informs us of the folding pathways. In section 3.2, we will discuss a simple thought experiment regarding such systems.

Section 4.2: Thought Experiment

Before addressing the results of multishape competitive anneals, it is worth developing the conceptual framework with which the anneals will be discussed. We have developed a manageable thought experiment system in which to address how a competitive anneal might progress. Figure 4.3 shows this system; it consists of two small nanostructures Origami A and Origami B. Each structure consists of six staple strands which binds an equal number of scaffold bases. The only difference between these structures is the location of the scaffold start/stop point. This is to say that the origin of the scaffold sequence bisects staple A1 for origami A, but falls between staples B3 & B4 for origami B. The binding pads of the two circle maps are quite similar, and can be interpreted as a simple 90° rotation.

Both structures are nearly identical, the staple binding pads were depicted as an outer arc for origami A, and an inner arc for origami B. As described in section 1.4, the black circle

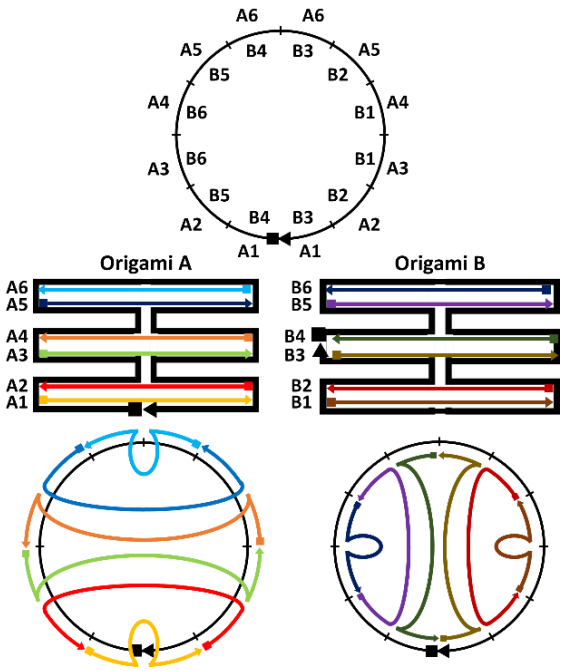


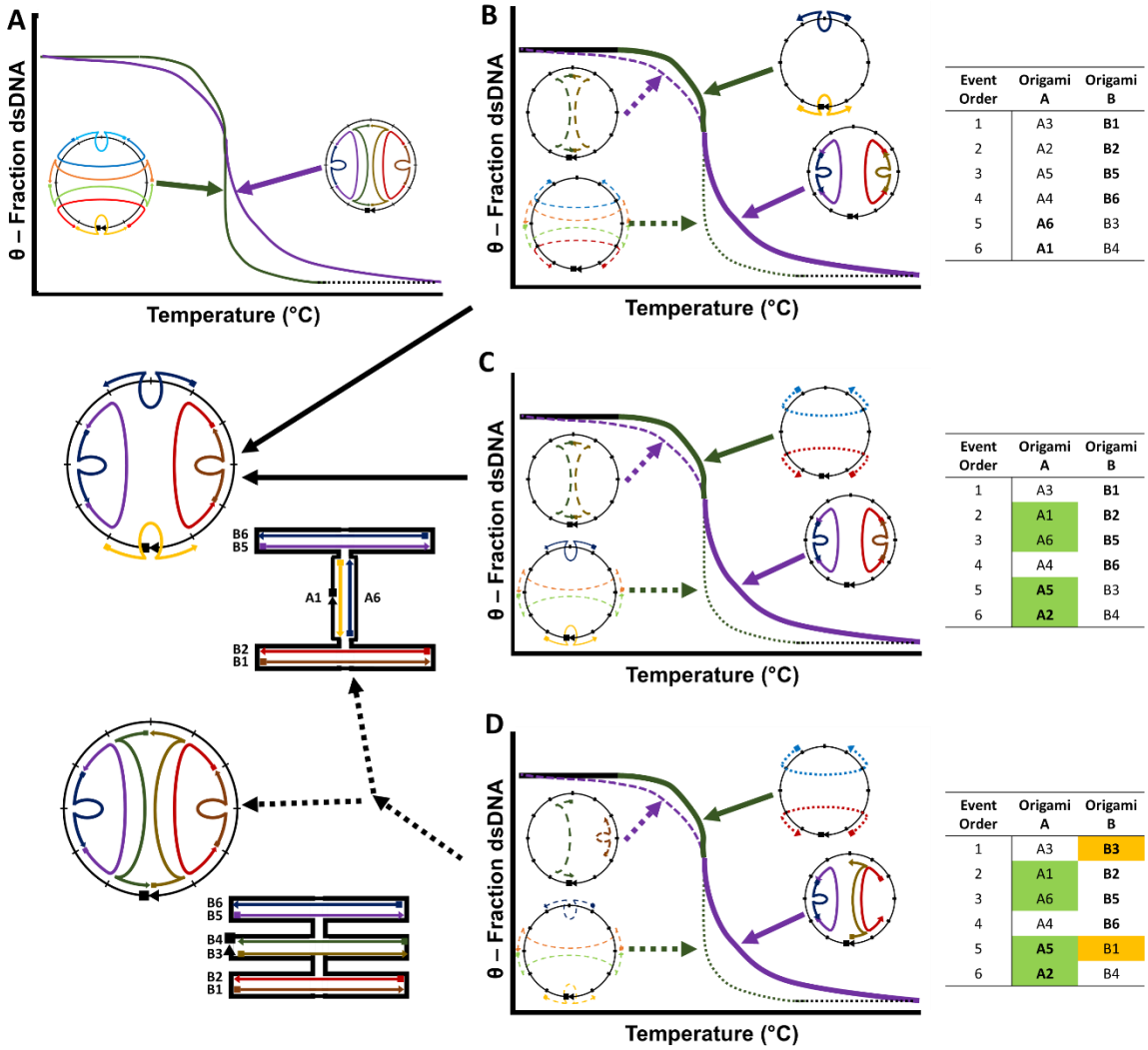
Figure 4.3: Small origami A and B: Top, map of competing binding pads; Middle, routing maps; Bottom, circle maps

represents the scaffold, where the sequence begins at the square and progresses to the triangle. The top section of Figure 4.3 shows the which staples in origami A and origami B compete for which sections of the identical scaffold.

Putting these theoretical nanostructures in competition is aided by the idealized annealing data in Fig 3.1. One limitation of annealing and melting data, whether it is obtained by DSC, UV-VIS, RT-PCR or some other technique, is the general inability to determine the order of distinct events in the progression between 0% and 100% formation. This is to say annealing data comes from ensemble techniques. An annealing curve only tells that the dsDNA content has increased, not which sections of a structure have formed. This limitation becomes significant when addressing competitive anneals. The idealized annealing curves for origami A and B are shown in Figure 4.4 A, as a green and purple line respectively. As these curves intersect at approximately 75% formation, the sections of the curve which have a higher formation % for a given temperature will be described as dominant. Those with the

lower % will be described as recessive. The temperature a reversible reaction occurs at is a function of the free energy of transition; An event with a higher annealing temperature should have a lower free energy of transition than a competing event with a lower temperature of transition.

As the annealing curves do not indicate which staples bind, and we are treating the formation as a path of dependent events, an unmanageable number of paths exists even for these simple systems. Without considering each staple binding pad separately, there are 6!,



Event Order	Origami A	Origami B
1	A3	**B1**
2	A2	**B2**
3	A5	**B5**
4	A4	**B6**
5	**A6**	B3
6	**A1**	B4

Event Order	Origami A	Origami B
1	A3	**B1**
2	A1	**B2**
3	A6	**B5**
4	A4	**B6**
5	**A5**	B3
6	**A2**	B4

Event Order	Origami A	Origami B
1	A3	**B3**
2	A1	**B2**
3	A6	**B5**
4	A4	**B6**
5	**A5**	**B1**
6	**A2**	B4

Figure 4.4: Hypothetical anneal data, staple binding orders, and theoretical chimera

or 720, possible binding pathways that a single of these small structures could follow. This complexity is increased further when competition is induced.

In one such set of paths, shown in Figure 4.4 B, the order of binding ensures that the nth event of each structure competes with the nth of the other. Such folding paths should form a single, well defined structure which inherits from its parent structures in proportion to the intersection of their anneal curves. As this structure is the combination of sections of two other structures, it will be called a chimera after monster of Greek mythology. In Figure 4.4 A, staples B1, B2, B5 & B6 all hybridize at a much higher temperature than their counterpart staples in origami A. At 75% formation, thee A1 and A6 staples bind, as they are more energetically favorable than their counterparts on origami B. The resulting chimera is shown below the annealing curve.

However, this first projected annealing process is not the only possibility. It is, however, more likely than a random ordering of which staples bind first. This is because competing staples share identical binding pads on the scaffold sequence. This implies an identical GC content and, as described in section 1.3, GC content is a powerful driver of hybridization. As such, staples with competing binding pads are likely to bind at similar temperatures and it is likely that the cooperative effects, conformational entropy, and the GC content of the staples' other binding pads will act as a tie breaker in determining which staple binds.

With that under consideration, even slight variations in the imagined pathway for these structures can make the system significantly more difficult to predict. Figure 4.4 B shows identical idealized annealing curves. For this case, the folding pathway of the origami B staples is identical, but the last three events of origami A are mixed. As one can see, the staple binding events which occur on the dominant curve compete with each other, and some of the events between the dominant and recessive curve are not competitive. This can be

seen as A1 and A6, which bind in the high temperature recessive section of the origami A annealing curve, would still bind before their competitors B3 and B4. One would anticipate this to result in the same chimera shape as in Figure 4.4 A, except through a less straightforward path.

Figure 4.4 C, presents an even less straightforward path. In this case, the path for Origami B is modified as well. In this case, a minimum free energy structure can only be found if either staples A1 and A6 displace staple B3, or if staples A1 and A6 remain half bound until being displaced by staple B4. The former will result in a chimera and the later in pure origami B.

These paths are by no means static or final, as discussed in section 1.3, strand displacement reactions can occur well below the transition temperature of a particular staple[14]. How such a set of reactions might occur for a competing pair of staples is shown in Figure

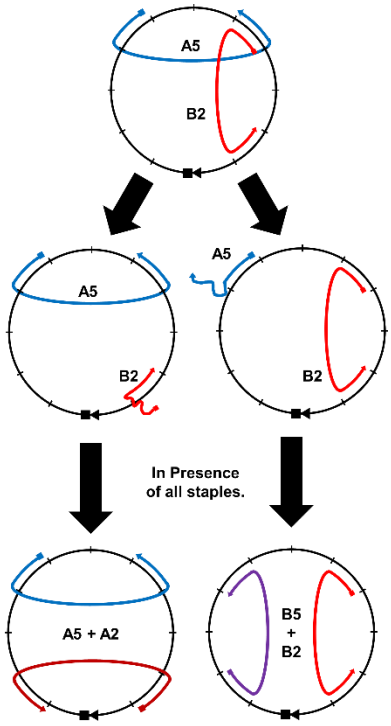


Figure 4.5: Strand displacement in the context of competitive anneals

4.5. The free energy of binding for the competing binding pads based on GC content is identical, so the path along which the strand displacements occur will likely depend on other factors. Such factors include the GC content at the staples' other binding pads, reductions in conformational entropy, and base stacking if the binding pads do not perfectly overlap.

Figure 4 shows the circle maps of six stable chimeras which could form in this idealized system. The small number of stable cases are limited due to the two-fold symmetry of the structures and the 90° of their circle maps. The staples of these structures compete in pairs; a stable structure with only one scaffold strand consisting of origami A and B staples can exist in only if either the set of staples A1 and A6 are bound or the set of staples B3 and B4 are bound.

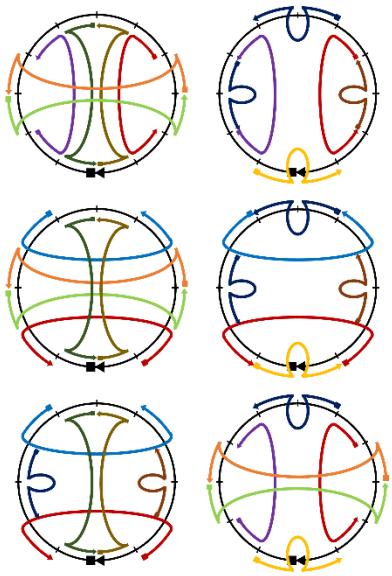


Figure 4 6: Stable chimera for theoretical origami A and B

Some of these chimeras may be physically impossible due to the physical constraints of DNA, such as the top left. Ultimately, it is highly unlikely that many radically different chimera will develop without complications. In particular, systems which can form multiple chimera types would also develop multiple scaffold chimera in which staples would bridge multiple

scaffold molecules to satisfy themselves. Such structures would likely polymerize, ultimately aggregating into large, chaotic messes which would be difficult or impossible to characterize.

Even if multiple stable chimera, a competitive system could still form a single set of rational, non-polymerizing, structures if strand displacement occurred along a path which favored a particular chimera. A competitive system which forms one, or few, chimera would be valuable as a window through which to examine the forces which drive annealing and strand displacement.

In this section we have described a simple system as a lens through which to understand more complicated competitive systems. In section 3.3, we discuss our initial experiments to confirm that rational nanostructures can form from competitive anneals.

Section 4.3: Initial Multishape System of Circle and Bundle

The circle and 6-helix bundle nanostructures were chosen as the initial test system for competitive anneals. One benefit of putting these structures into competition is the sheer difference in their routing and circle maps, as shown in Figure 4.7. This stark difference between where their staples bind should ensure that a section which begins to form one shape will not benefit from cooperative effects. As the 6 helix bundle is a 3D structure, it allows for height contrast which immediately distinguishes it.

To maximize the probability of observing chimera nanostructures, the systems were annealed at relatively high concentration, 25 nM and 50 nM scaffold and at a relatively low staple excesses. All samples were annealed in TAE buffer with 12.5 mM Mg, over a period of 8 hours from 80°C to 25°C. For each concentration, three samples were annealed with 5x total excess staple, with ratios of circle/bundle of 1:2, 1:1, and 2:1. Example AFM images of these samples are shown in Figure 4 A-F. Figure 4.7 A-C correspond to 50 nM scaffold, and D-F correspond to 25 nM. The staple excess described above correspond to A&D, B&E, and

C&F respectively. It is worth noting obtaining a truly accurate evaluation of a sample requires more than the 7-40 nanostructures which are easily visible in a 1-2 μm scan.

From Figure 4.7, the differences in structure types, and in overall character of the images are clear. We define rational chimera as nanostructures in which the species of each section is distinguishable, and which do not excessively polymerize multiple scaffold molecules. Rational chimeras are clearly visible in Figure 4.7. Some stable multiscaffold chimera as in Figure 4.7 F, are rational and do not polymerize in to large aggregate structures, other multiscaffold chimera, Figure 4.7 B and E, do polymerize to form large clumps which are difficult to characterize. The quality of these images is of some concern, as the number of interconnected aggregates makes characterization difficult. The presence of large aggregates, which are cropped from these images, creates difficult to remove flattening effects as in Figure 4.7 E.

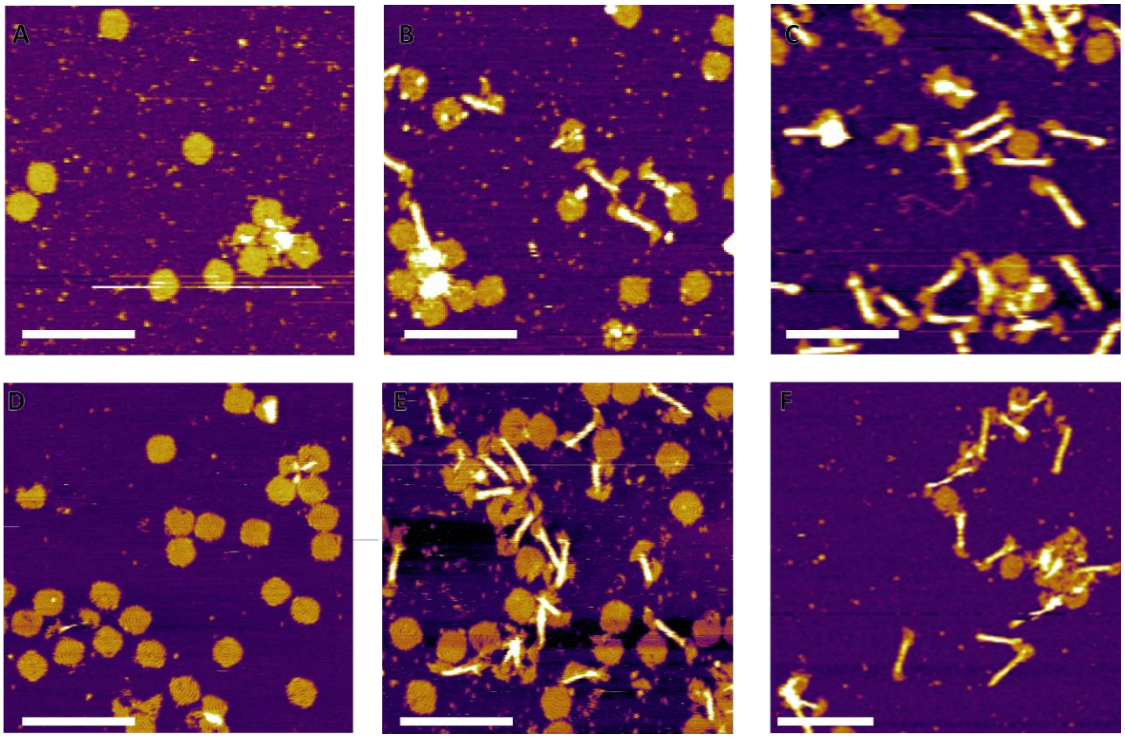


Figure 4.7: Initial 5x staple excess anneals: A-C: 50 nM; D-F: 25 nM; A, D- 3.33xCircle/1.66xBundle; B,E- 2.5xCircle/2.5xBundle; C,F- 1.66xCircle/3.33xBundle. Scale bars 200 nm

From these samples, features of interest were cropped and emphasized in Figure 4.8 A, E, and F are of interest as they clearly show the chimera with distinct transitions between shapes. Figure 4.8 B, D, I and J all show interesting aggregate behavior and improper anneal behavior in which chimera may be identified, but not easily analyzed. Finally, Figure 4. H is of interest as it depicts a particularly clear and rational multiscaffold chimera.

After the initial samples were imaged, further Triangle/Bundle and Triangle/Circle samples were made as per the circle/bundle above. All of which formed pure triangle structures, shown in Appendix 5. As a search for chimera forming conditions could become exhaustive, a final set of ternary samples was tested to examine the degree to which the triangle was able to dominate the other shapes. These systems were annealed as above, with 25 nM scaffold and an overall staple excess of 7.5x. This excess was divided so that bundle and circle staples were at a 1:1 ratio, with decreasing excesses of triangle ranging from 1/3, to 1/5, to 1/20 of the total excess. This is shown visually on a ternary diagram in Figure 4.9. These ratios meant that no individual shape had more than 5x excess staples, the commonly accepted threshold for acceptable yield, with the triangle having as low as a 0.5x excess.

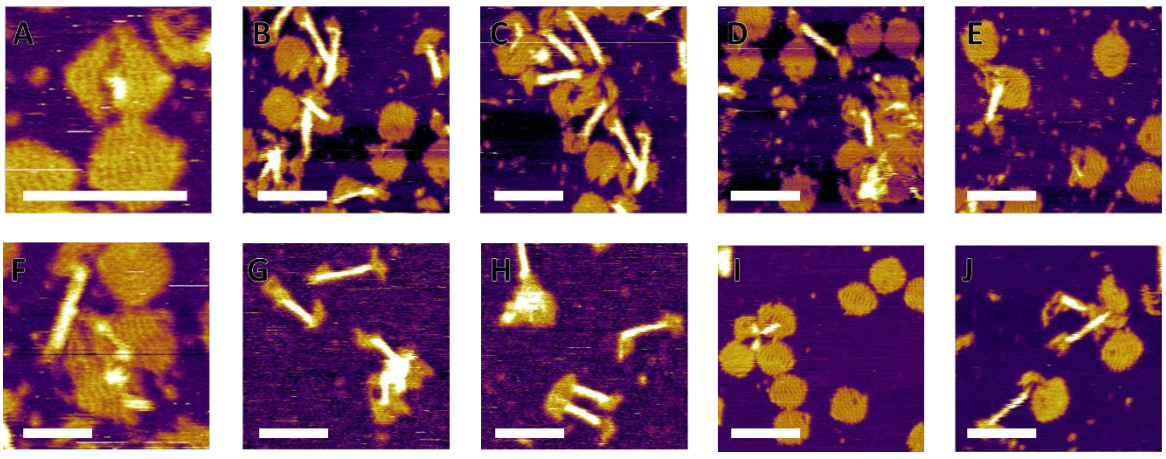


Figure 4.8: Close up images of chimera nanostructures. A-F 25 nM scaffold 1.66xCircle/3.33xBundle; G,H 50 nM scaffold 2.5xCircle/2.5xBundle; I,J 25 nM scaffold, 1.66xCircle/3.33xBundle. Scale bars 100 nm

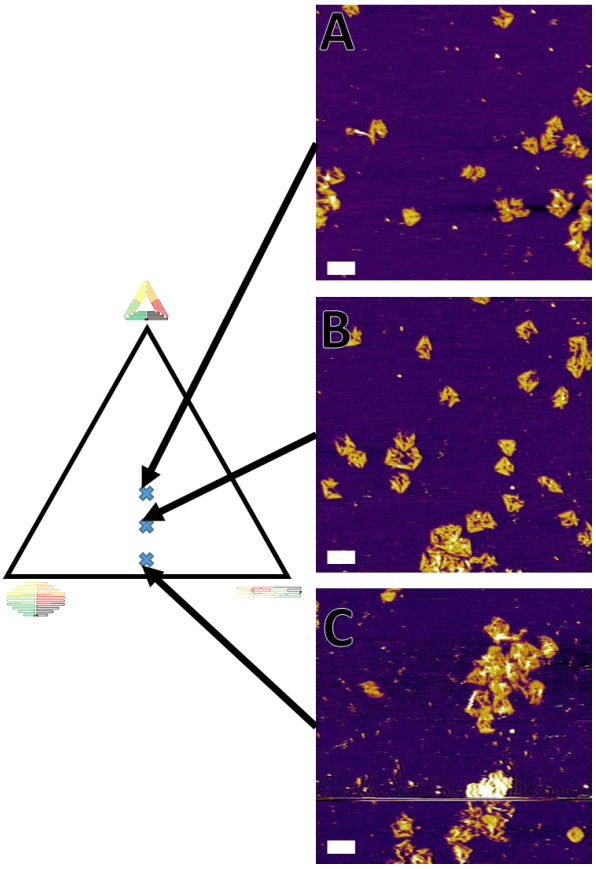


Figure 4.9: Ternary competitive anneals 7.5x total staple excess. A- 2.5xTriangle/2.5xCircle/2.5xBundle, B- 1.5xTriangle/3xCircle/3xBundle, C-0.5xTriangle/3.5xCircle/3.5xBundle. Scale bars 100 nm

As shown in Figure 4.9 A-C the triangle shows surprising dominance. At 2.5xTriangle/2.5xBundle/2.5xCircle, Figure 4.9 A, the scaffold nearly completely displays triangle structure, albeit with a slightly higher rate of open seams than in pure triangle anneals. At 1.5x, Figure 4.9 B, the triangle is still clearly dominant, but is rarely seen fully formed and without defect. As shown in Appendix 6, this sample showed significantly sized aggregate particles. Finally, at 0.5x excess, Figure 4.9 C, sections of the triangle can still be seen, but large malformed aggregates represent most of the AFM scan.

These samples suggest strand displacement cannot fully push an arbitrary system to stable chimeras at low staple concentration. Higher excesses should encourage displacement. Given the presence of the large aggregates in the 0.5x Triangle sample, it is

likely that some number of the triangle staples bound at high temperature to the scaffold and were never displaced, resulting in polymerization of scaffold molecules by unsatisfied bundle and circle staples. In the context of origami design, scaffold binding pads are designed to take up similar lengths of bases for all structures, so it is possible that some triangle staples benefitted from cooperative effects of the other shapes. This is to say a staple may not need to be from an identical structure as its neighbor to benefit from cooperative factors. Another point of relevance is that the triangle, so minimally reduces conformational entropy at each step. This can be seen in the triangle staple map in Figure 4.9 A; the triangle has few to no staples which bind truly disparate scaffold sections.

As seen in Figure 4.7, Figure 4.8, and Figure 4.9, we have clearly confirmed the creation of chimeric DNA nanostructures. These structures hint at the potential of competitive multishape anneals to produce complex and interesting structures without additional investment in design or synthetic DNA. However, as is also evident by the damaged structures and the presence of large aggregates, these systems are suboptimal for characterization.

Section 4.4: Binary System Characterization

Given the complexity of the images shown in Section 3.3, it was decided that a change in approach was necessary. The total staple excess was raised to, and kept constant at 7.5x. Only a relatively short range of excesses 2.5x/5x, 3.75x/3.75x, and 5x/2.5x were examined. These samples were annealed in TAE Mg, at scaffold concentrations of 5 nM, 25 nM, and 50 nM, over 2 hr, 8 hr, and 16hr. This testing grid, described in Table 4.1 resulted in a set of 27 samples. These samples were imaged until a 50-200 structures of acceptable quality were found, approximately 2-8 images.

Table 4.1: Sample grid for competitive anneal systems

Staple Ratio	**Concentration**	**Time/Rate**
2.5x Circle 5.0x Bundle	5 nM	2 hr 0.58 °C/min
3.75x Circle 3.75x Bundle	25 nM	8 hr 0.15 °C/min
5.0x Circle 2.5x Bundle	50 nM	16 hr 0.07°C/min

These nanostructures were counted manually, with random confirmation checks, and sorted into categories of pure circle, pure bundle and chimera. After counting was performed, the chimera and bundle samples were analyzed to determine the portion of the chimera which were in either the bundle or circle shape. To do so the height difference between the origami was utilized. As in Figure 4.10, the images were entered into imageJ[15], and the image was thresholded to only select bundles, with aggregates blacked out before analysis. The thresholded images were then subject to particle analysis and fitting to an oval/eclipse. The major axis of the oval/eclipse was determined to be “length” of the bundle portion of the chimera. As shown in the appendix 7, the minor axis of the fitted ovals was remarkably consistent at 18 nm. The overlay map was then used to check that only the fitted ovals which matched a chimera were used, and the total number of chimera from particle analysis was compared to the counted number.

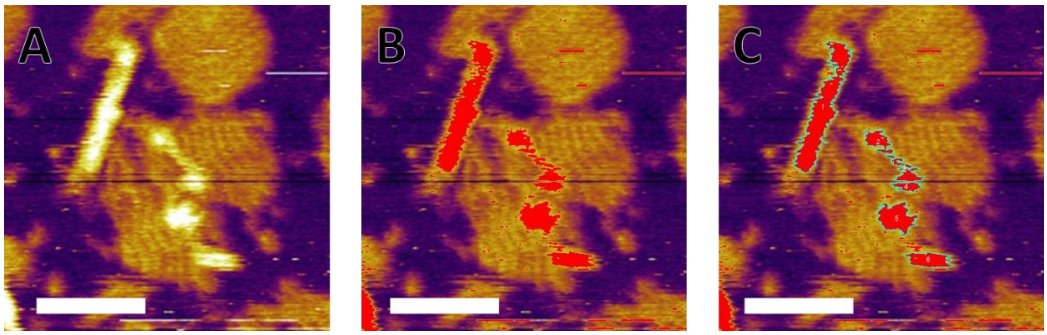
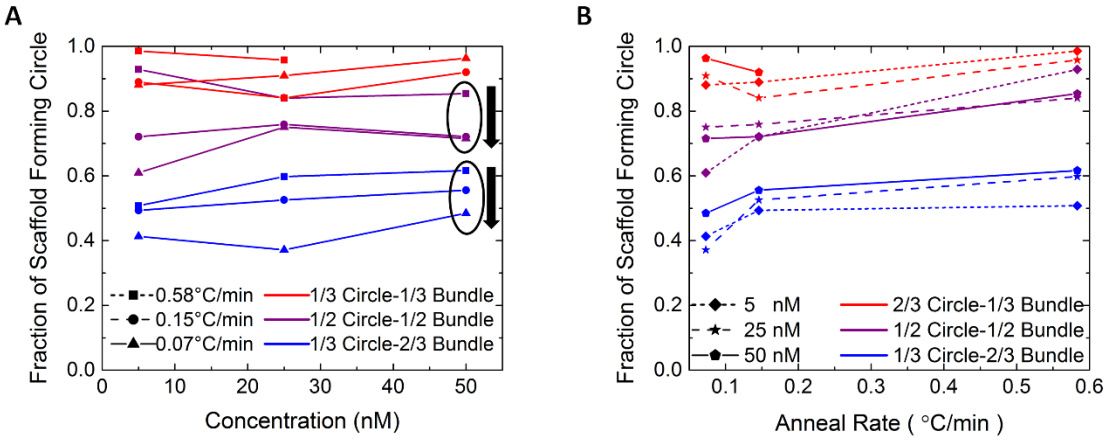


Figure 4.10: Image analysis for Chimera Structures. A-Raw image, B-Thresholded image, C-Particle analysis of thresholded image. scale bars 100 nm

After applying the correction factor, these values were used, in conjunction with the maximum length of pure bundle measured, 138 nm, to determine the fraction of each chimera that was bundle and circle respectively. This was done by dividing each chimera's bundle length by the maximum bundle length and assuming the remainder of the scaffold formed circle.

Figure 4. shows two plots of the resulting data. Given inherent difficult in gathering statistical data from AFM, it is expected that there is some significant level of noise. Figure 4.11 A, plots the fraction of scaffold forming circle as a function of concentration, with each line addressing a single staple ratio and anneal rate. While there is some variability, the fraction of each sample forming circle is seen to decrease with a slower anneal rate, particularly for samples with a higher excess of bundle staples. The differences between annealing rates pale in comparison to those between staple excesses, but are visible.

As shown in Figure 4.11 A, the curves for lower annealing rates tend to exhibit a slightly higher fraction of scaffold forming the circle shape. Additionally, there appears to be a slight



reduction in circle formation as concentration increases. These differences are most noticeable for the blue curves which have twice as much bundle staple excess to circle staple

Figure 4.11: Length corrected data for fraction of scaffold in each sample forming circle. A-as a function of anneal concentration, B-as a function of annealing rate

excess. The trends in Figure 4.11 A are mirrored in Figure 4.11 B, which replots the same data as a function of annealing rate.

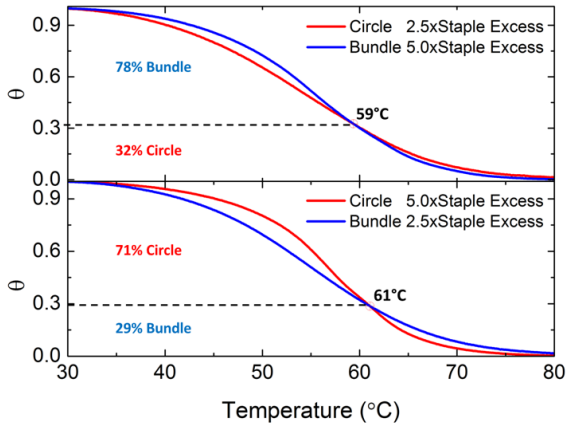


Figure 4.12: Fluorescence anneal curves and intersections for A-2.5xCircle/5xBundle, B-5xCircle/2.5xBundle. All 4 curves were annealed separately

In order to place the fraction forming each structure in the context of annealing processes, it was compared to fluorescence spectroscopy data, gathered and processed as per Chapter 3. This data was only gathered for staple excesses in multiples of five and for 5 nM and 15 nM concentrations. As such the best points of comparison for the AFM gathered data is the anneal curves for 5 nM Bundle and Circle at 2.5x and 5x staples. These curves are plotted against each other in Figure 4.12 where the upper plot shows an excesses of 2.5xCirlce/5xBundle and the lower shows a 5xCircle/2.5xBundle. In both plots a clear intersection occurs at 28-32% formation. In both cases the curve associated with the higher staple excess is dominant for a larger percentage of formation.

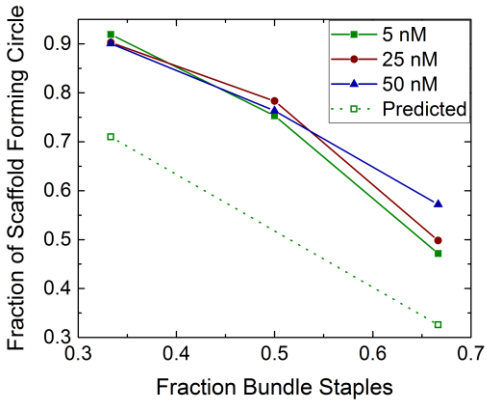


Figure 4.13: Averaged fraction of scaffold as a function of fraction bundle staples, compared to intersections of anneal data

Figure 4.13 shows the data from Figure 4.11, averaged over the annealing time, compared to the percentage circle which one would predict from the annealing data shown in Figure 4.12. It is of import to note that the curves have a very similar slope to the predicted points, and all of the concentrations fall on a similar curve. A source of error also worth noting is the correction by length of the bundle section of the chimera. The bundle length 138 nm was taken from the maximum length seen for a bundle. However, if one were to average a bundle sample set, neglecting poorly formed structures, the average length would be approximately 107 nm. This difference is enough to shift the fraction of scaffold forming chimera down by ~0.1, or half of the distance between the measured and predicted values.

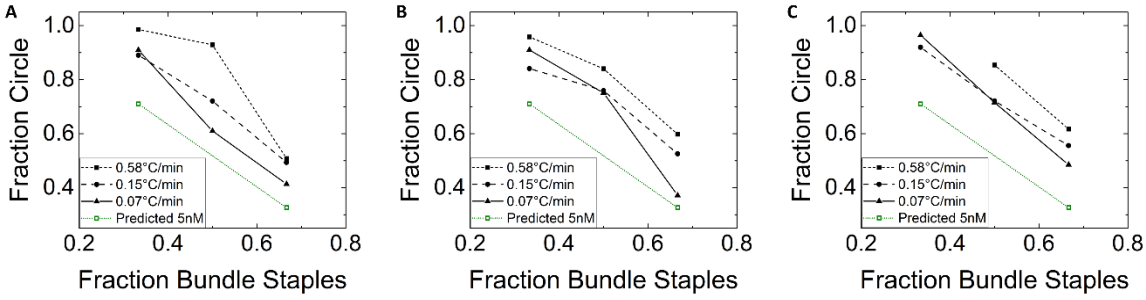


Figure 4.14: Fraction of scaffold forming circle as a function of fraction bundle staple, A-5 nM samples, B-25 nM samples, C-50 nM samples

As the average bundle length can change dramatically depending on what one qualifies as a poorly formed structure, it made sense to stick with the maximum on the understanding that the presented data will have some vertical shift.

To further explore the role of annealing rate on the structure formation, the data in Figure 4.14 was separated into three plots, Figure 4.14, by concentration, and compared to the predicted value as a function of the fraction of total excess staples which belong to the bundle shape. As is clearly visible, there is a general downward trend in fraction of scaffold forming circle as the annealing rate decreases. This strongly implicates the role of strand displacement in guiding the chimera toward their final structures. It is also worth noting that particularly slow anneal times are closer to the circle formation predicted by anneal curves.

In addition to the insight they provide for the overall condition of competitive multishape anneals, these curves also indicate that it is reasonable to use the annealing data as a benchmark for competitive systems. Given that the measured fraction of each system in circle shape could be shifted up or down based on the bundle length correction factor, it would be inappropriate to use the anneal data to predict how much of a given system is acting as a chimera, especially given the reality of strand displacement. However, it is reasonable to use anneal data to predict when a given competitive system will exclusively form a single shape, and from that when it is likely that chimeras will form. However, before addressing true prediction, the chimeras must be analyzed as structures.

In this section we have discussed quantitative analysis of the chimera nanostructures via AFM, our correction factor for the AFM data. We showed the relationship between the fraction of the scaffold molecules forming each shape, and how that related to the fraction of staples from each shape, the concentration and the anneal time. In section 3.5 we will discuss qualification of the nanostructure routing map from AFM data.

Section 4.5: Understanding Chimeric Structures with Qualitative AFM

To complement the data gathered in Section 3.4, we will now address the structure of competitive multishape chimera. To do so, we must provide a spectrum of the chimera formed. The AFM data will then be used to infer the routing of the scaffold molecule through the chimeras, which in turn will shed light on the formation of the chimera.

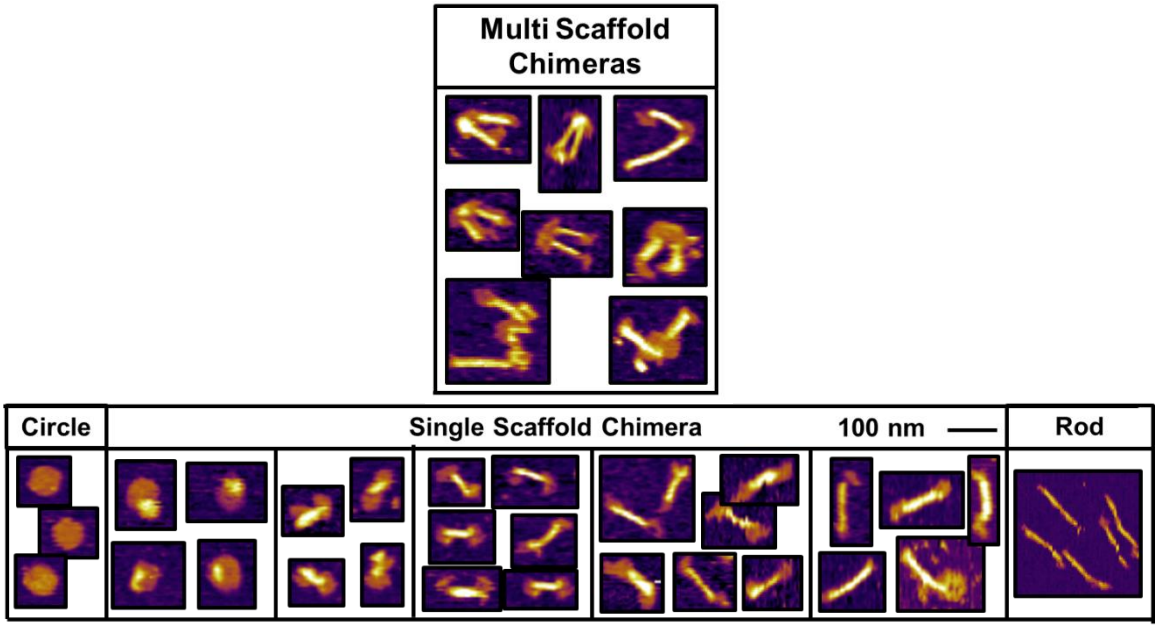


Figure 4.15: Individual chimera sorted by bundle section length and multiscaffold chimeras. Scale bar 100 nm

Example chimera were taken from all anneals, sorted, and displayed in Figure 4.15. The chimeras were sorted by the length of the bundle section of the chimera. Each sample contained a distribution of these chimera. Samples with a higher fraction of scaffold forming circle were more likely to have chimera with a higher circle content, but he chimera which appeared in each sample did so across a distribution of bundle segment lengths. This could be attributed to a gradual strand displacement process, or to the depletion of staple excesses during formation, or independent nucleation of the shapes followed by growth. In all cases the center of the bundle structure was dominant over the circle center, while the ends of the circle were dominant over the bundle ends.

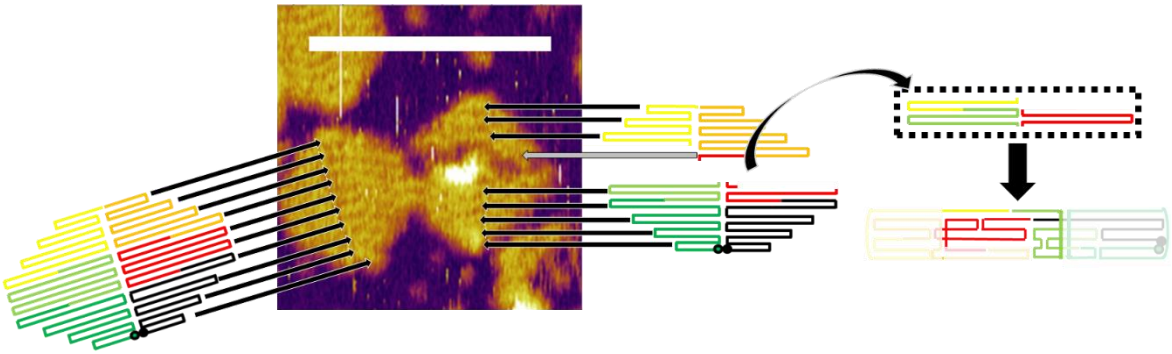


Figure 4.16: Image analysis of chimera origami and comparison to routing maps. 25 nM scaffold, 2.5xCircle/2.5xBundle, scale bar 100 nm

There implications of manipulating and forcing multiscaffold chimeras to form is particularly interesting for creation of much larger shapes. If the polymerization of shapes could be predicted and controlled, it would allow the creation of large structures which could potentially combine the addressability of tiled DNA nanostructures with the addressability of scaffolded structures. The multiscaffold chimeras were relatively rare, and were most prevalent in samples with large aggregates which are not conducive to AFM imaging, section 3.2. These factors ensured that there were fewer multiscaffold chimeras to be presented in Figure 4.15.

Of the AFM images gathered for the competitive circle/bundle system, several were of particularly high quality. One such, Figure 4.16, contained several features which allowed more in depth analysis of the chimera structure. In particular, the pure circle on the left of Figure 4.16, shows clearly defined ridges at the seam of the origami, which match perfectly to the number of seams in the routing map. This provides confidence in analysis of the chimera which is base stacked to the pure circle. Additionally, the bottom of each circle has a slight bulge, indicating the loop of excess scaffold ssDNA at the start/end of a structure's routing. By comparing the number of seam ridges in the chimera, to the same routing map, the section of the scaffold forming the bundle shape can be determined.

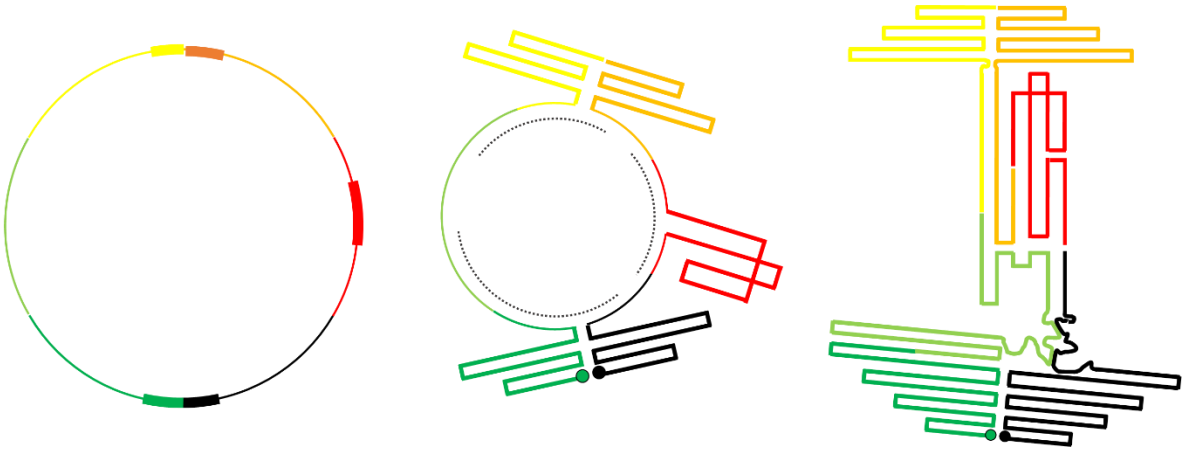


Figure 4.17: Nucleation and growth in chimera formation.

This information, along with the trends regarding strand displacement discussed in section 3.4, were used to develop the following observations regarding circle/bundle chimera. For any DNA scaffold, there will be sections of the scaffold with higher GC content, which will have more favorable energies of binding. For most structures, many staples exist whose reduction in conformational entropy is so great that they can only bind well after their neighbors. Therefore, for most structures, some of the high GC content sections will hybridize at a much lower temperature, than they otherwise would.

If two routing maps are different enough that they have different regions of conformational entropy destabilized high GC content, when brought into competition, each shape will nucleate at the disallowed region of the other. From these nucleation points, each shape will grow and compete by strand displacement, driven by the staple excess associated with that shape. This proposed process is generally described in Figure 4.17 and will be described as the nucleation and growth model of chimera formation.

Using this hypothesis, and the tools applied to single scaffold chimeras, a tentative, preliminary routing map for multiscaffold chimeras as created, shown in Figure 4.18. As the bulb is of identical width to a pure circle, and it is unlikely that three scaffolds were present,

the region in which the two scaffolds were bound together must occur in the conjoined bulb. In addition, the two scaffold molecules must have relative proximity and an unimpeded entrance/exit in order to exit the routing map of the circle and enter that of the bundle.

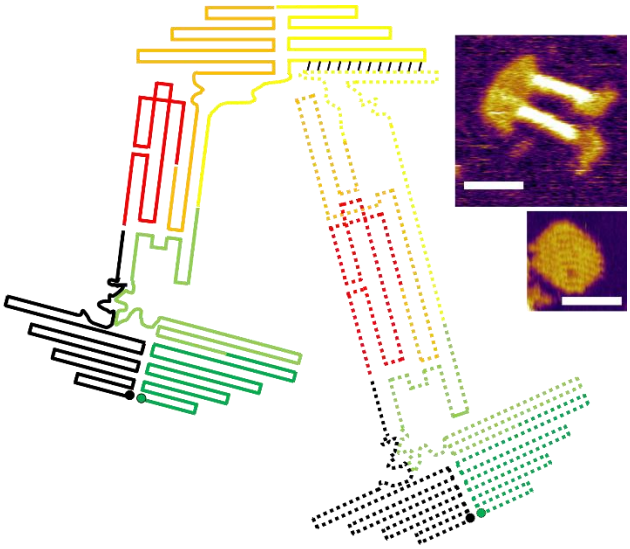


Figure 4.18: Tentative multiscaffold chimera routing map. Scale bars 50 nm

In this section, the AFM data provided insights into the routing of chimera nanostructures. This routing was used to propose the hypothesis that chimera formation is driven by nucleation and growth of shapes along the scaffold. This hypothesis indicates that the fraction of a chimera forms a shape is determined at least partially by the growth rate of that shape, which in turn is driven by staple excesses. Collectively these results explain the behavior of both the circle/bundle binary and circle/bundle/triangle ternary system. The triangle, with its lower reductions in conformational entropy, could feasibly nucleate at the same positions on the scaffold as both the circle and bundle, allowing its staples to poison the scaffold and prevent proper formation of the other shapes. In section 3.6, we will discuss use of anneal data, as described in chapter 3, to predict phase diagrams of the competitive systems.

Section 4.6: Prediction of Phase Diagrams via Anneal Data

As discussed in the previous section, 4.6, the nucleation/growth model of chimera development is proposed. The θ, or fraction of dsDNA, can be described as a measure of growth a structure. Given that anneal data shows the fraction of scaffold which will form a given structure, section 3.2, it can be used to predict some, although not all chimera growth as it does not predict strand displacement. Interpretation of the curves was performed by assuming that shapes which had a significantly higher θ over all temperatures would make completely formed nanostructures of that shape. Due to strand displacement, the character of intersecting curves is difficult to predict.

By differentiating pairs of curves for which one structure is dominant from those which intersect it is possible to generate phase diagrams for competitive multishape DNA nanostructures using anneal curves for the independent shapes. This is exceptionally powerful not only because it leverages the high throughput capacity of RT-PCR equipment, but also because it allows a single anneal curve to be used in comparison to multiple other curves. This maximizes the data garnered as a function of experimental time. Particularly given the time consuming nature of AFM, this is a valuable capability.

For the purpose of developing phase diagrams, we chose to use staple excess, rather than overall concentration, as an axis of interest. This choice was made partially due to the relative quantities of scaffold and staples which were in stock and partially due to the relative consistency of signal/noise of RT-PCR fluorescence data as a function of staple excess, discussed in section 3.3. Anneal data for the three multishape structures, circle, bundle, and triangle, were gathered at seven different staple excesses with three replicates of each sample. The anneals were performed in Na-Cacodylate buffer as its pH is stable as a function of temperature and with the standard 1 mM EDTA/12.5 mM Mg. The staple excesses were

2.5x, 5x, 7.5x, 10x, 15x, 20x, and 40x. All anneals discussed in this section occurred at 5 nM scaffold. The anneal curves, after data processing discussed in section 3, is shown in Figure 4.19.

The anneal curves in Figure 4.19 indicate that the triangle anneals at a higher temperature than either the circle or bundle. These curves also indicate that the circle has a higher variation in the breadth of the annealing transition than either the bundle or triangle, but is also anneals at a slightly higher temperature than the bundle.

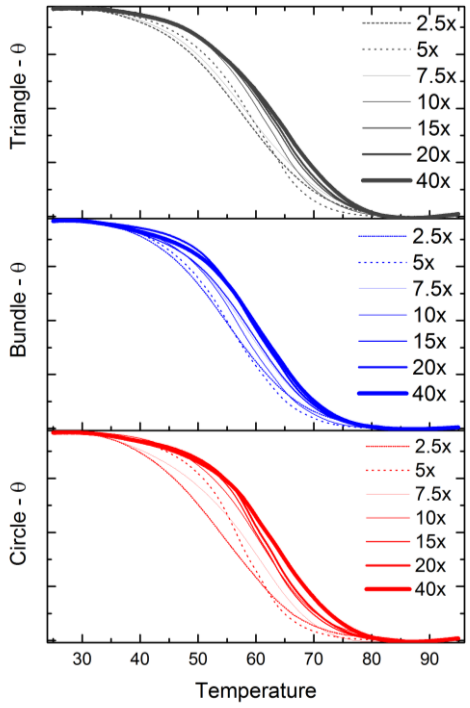


Figure 4.19: Fluorescence anneal curves for Triangle, Bundle and Circle miniM13 structures at a variety of staple excesses

Curves of different shapes were compared, given a classification for predicted 'phase', and plotted as a function of total staple excess and fraction of staples from each shape. Curves which intersected anywhere between 10-90% formation were classified as chimera/unknown and are depicted as crosses. This window was chosen as the events

happening below/above 10/90% completion would likely be single binding pads which could be easily strand displaced. Anneal curve pairs in which one shape is dominant were signified by red circles, blue squares and grey triangles for the circle, bundle and triangle shapes respectively, shown in Figure 4.20 A-C. As is visible in raw anneal data, the triangle has the strongest driving force of the three, with the circle having a slightly stronger driving force than the bundle.

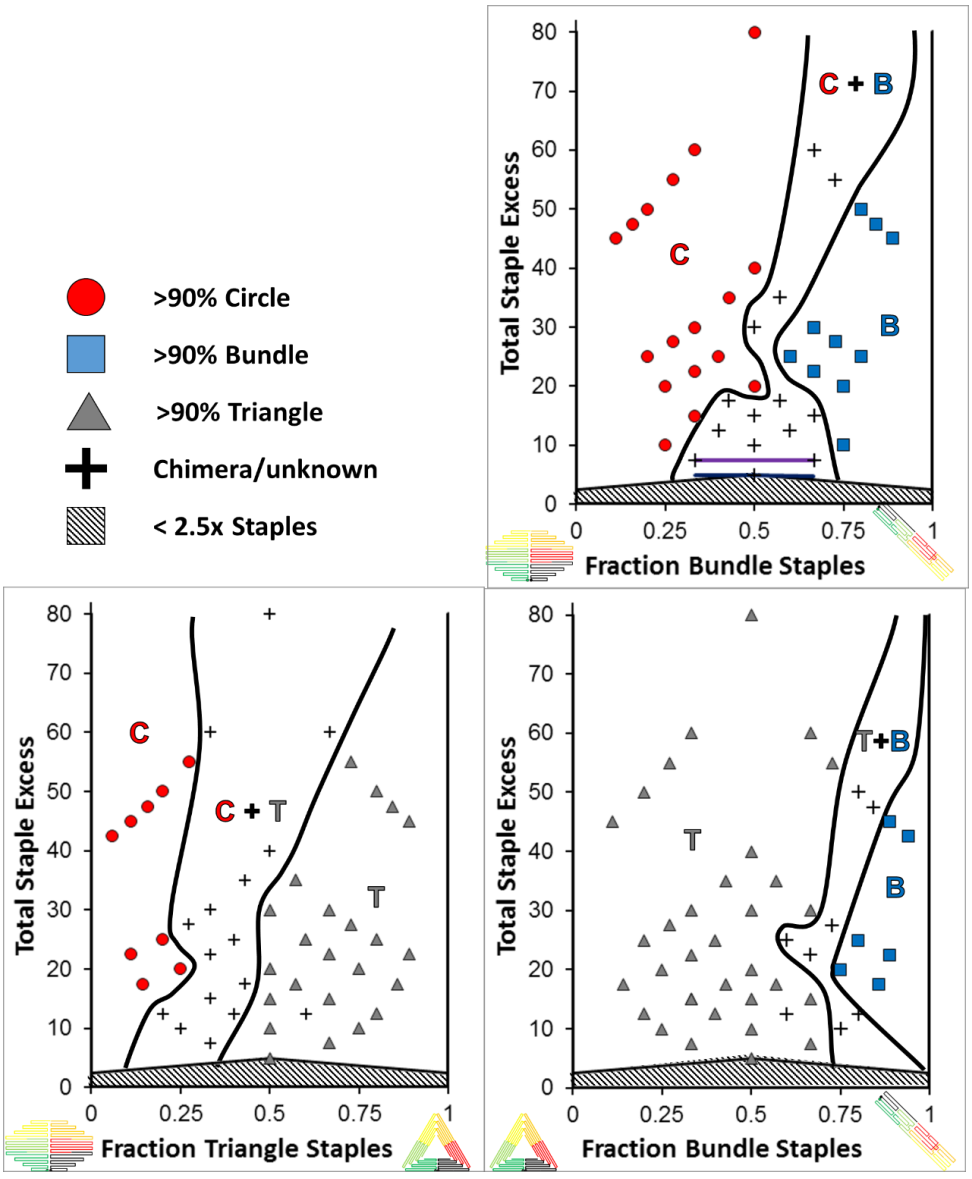


Figure 4.20: Predicted binary diagrams for circle/bundle, circle/triangle, and triangle/bundle. Purple and blue lines in circle/bundle diagram indicate experimental data from sections 4.3 and 4.4 respectively

All three diagrams have a change in width of the chimera/unknown region between 20-40x total staple excess. This is reminiscent of hysteresis data shown in section 3.5 Figure 3.15. The hysteresis area is expected to be correlated to cooperativity. This area was only gatherable for the Tall Rectangle M13 structure, whose anneal data was much cleaner and more consistent than the smaller miniM13 structures. The hysteresis area for the Tall Rectangle had a similar and interesting trend for both staple excess and scaffold concentration, which after a minimum proceeded to increase. This implies that some aspect of the cooperativity is dependent on staple excess and scaffold concentration. As cooperativity plays a major role in the annealing and melting of these structures, one would expect such a curvature in the cooperativity to significantly change the interplay between the annealing of two structures. However, further experiments are required to determine whether this is a true trend, or some artifact of photobleaching or other human error.

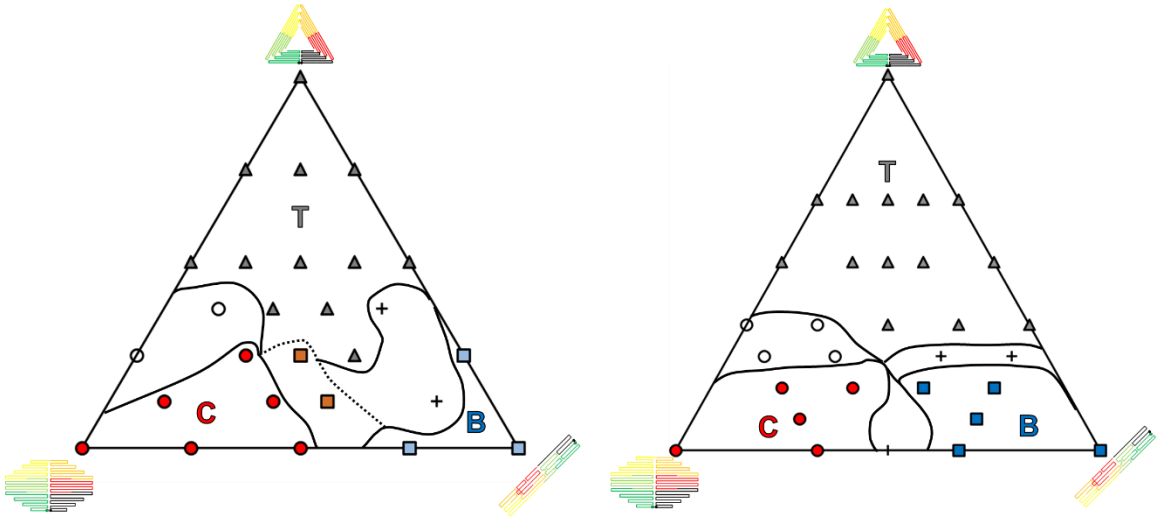


Figure 4.21: Predicted ternary phase diagrams from anneal data for A- 20x total staple excess, B- 30x total staple excess

Similar to the production of binary diagrams, two ternary diagrams were produced for 20x and 30x total staple excess, as shown in Figure 4.21. The brown squares indicate compositions for which all three shapes participated in the dominant curve. Given the number

of curve combinations which add to a total of 20x and 30x respectively, the diagrams have a different number of data points. In comparison between the diagrams, it is reasonable to interpret that the chimera/unknown regions expand at lower concentrations, as in the binary diagrams. However, the degree to which this case is difficult to assess given the different number and position of data points. This expansion of chimera regions as a function of decreasing staple excess likely explains the poorly formed ternary anneals at 7.5x total staple seen in Figure 4.9.

In this section, the anneal data and the nucleation/growth hypothesis for chimera formation, were used to predict the phase/shape of a competetive multishape anneal. The potential for these predictive capabilities was demonstrated. However, true implementation of these capabilities requires both higher reliablity of anneal data via photobleaching corrections and AFM testing of the composition in the phase diagrams.

Section 4.7: Conclusions and Future Work

In this Chapter, we have discussed the utility of new techniques designed to probe nanostructure formation, section 4.1; created a thought experiment system to address the complexity of competitive multishape anneals, section 4.2; confirmed the ability of nanostructures to form despite competitive conditions and confirmed the formation of chimera nanostructures, section 4.3; performed initial analysis on select competetive anneals evidencing the occurance of strand displacement and the qualitative predictions of anneal data, section 4.4; examined the structures of individual chimera to develop a hypothesis behind their formation, section 4.5; implemented this hypothesis, and the spectroscopy techniques discussed in Chap. 3, to generate predictive phase diagrams for competetive anneals, section 4.6.

While the work discussed here provides an initial step in probing the complicated dynamics of a competetive anneal, further analysis must be performed. In particular, future work should focus on testing the predictive abilities of anneal data, further probing the role of strand displacement in nucleation/growth, as well as performing similar analysis on larger M13 origami as conformational entropy will play a stronger role for larger scaffolds. Once these basic parameters are understood, future work will be directed to designing scaffolded DNA nanostructures with predetermined chimera as well as structures whose competition is designed to probe specific coopertivity/conformational entropy relationships.

References

1. Dannenberg, F., Dunn, K. E., Bath, J., Turberfield, A. J. & Ouldridge, T. E. Modelling DNA Origami Self-Assembly at the Domain Level. 1–22

2. Arbona, J.-M., Aimé, J.-P. & Elezgaray, J. Cooperativity in the annealing of DNA origamis. *J. Chem. Phys.* **138,** 015105 (2013).

3. Arbona, J. M., Elezgaray, J. & Aimé, J. P. Modelling the folding of DNA origami. *EPL (Europhysics Lett.* **100,** 28006 (2012).

4. Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459,** 414–8 (2009).

5. Song, J. *et al.* Direct visualization of transient thermal response of a DNA origami. *J. Am. Chem. Soc.* **134,** 9844–7 (2012).

6. Högberg, B., Liedl, T. & Shih, W. M. Folding DNA origami from a double-stranded source of scaffold. *J. Am. Chem. Soc.* **131,** 9154–5 (2009).

7. Estevez-torres, A. Observing and Controlling the Folding Pathway of DNA Origami at the Nanoscale. (2016). doi:10.1021/acsnano.5b05972

8. Song, J. *et al.* Isothermal hybridization kinetics of DNA assembly of two-dimensional DNA origami. *Small* **9,** 2954–2959 (2013).

9. Dunn, K. E. *et al.* Guiding the folding pathway of DNA origami. *Nature* (2015). doi:10.1038/nature14860

10. Wei, X., Nangreave, J., Jiang, S., Yan, H. & Liu, Y. Mapping the thermal behavior of DNA origami nanostructures. *J. Am. Chem. Soc.* **135,** 6165–6176 (2013).

11. Sobczak, J.-P. J., Martin, T. G., Gerling, T. & Dietz, H. Rapid folding of DNA into nanoscale shapes at constant temperature. *Science* **338,** 1458–61 (2012).

12. Fern, J., Lu, J. & Schulman, R. The Energy Landscape for the Self-Assembly of a Two-Dimensional DNA Origami Complex. *ACS Nano* acsnano.5b05309 (2016). doi:10.1021/acsnano.5b05309

13. Whitelam, S., Schulman, R. & Hedges, L. Self-assembly of multicomponent structures in and out of equilibrium. *Phys. Rev. Lett.* **109,** (2012).

14. Zhang, D. Y. & Seelig, G. Dynamic DNA nanotechnology using strand- displacement reactions. *Nat. Publ. Gr.* **3,** 103–113 (2011).

15. Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with imageJ. *Biophotonics International* **11,** 36–41 (2004).

Chapter 5: Chemical Quenching to Directly Monitor Folding Pathways

Section 5.1: Introduction and Motivation

The previous chapters of this work have addressed the folding of nanostructures through indirect methods, leaving direct techniques to this chapter. This chapter will present a novel method for chemically quenching partially formed structures along the folding pathway and stabilizing them for analysis by AFM. Before discussing such examination of folding pathways, its importance and background will be discussed.

A robust experimental platform which examines the folding of DNA nanostructures as it occurs would be intensely valuable for DNA nanotechnology as a field. Just in the work presented for chapters 2-4, such a platform would provide key additional information. For example, the role of scaffold length in folding could be probed by comparing miniM13 and M13 structures. Appropriate quench data for a sample could provide context for specific details in anneal curves. Examination of the folding processes during competition, particularly for knock out experiments in which parts of different structures are omitted, would provide key understandings necessary to design chimera nanostructure systems.

Beyond the context of the work presented in this thesis, such an experimental platform would prove an invaluable complement to numerous literature works. The effect of solution conditions on folding order could be determined and designed; Buffer solutions could be tuned to increase or decrease driving force contributions from hydrophilic, hydrogen bonding, and hydrophobic base stacking interactions[1–3]. Dynamic structural systems could be frozen and examined directly, allowing examination of intermediate states[4–8]. The folding for isothermal and chemical anneals could be examined and compared to typical thermal ramp systems[9,10].

The value proposition described above for experimental platforms which explore folding are merely extensions of work already done by others and ourselves. Such a platform would surely enable new, more valuable avenues of research. However, creation of such a platform has obstacles, which we will discuss in the section 5.2.

Section 5.2: Previous Attempts to Monitor Folding

Given the importance of folding process to DNA nanotechnology, an impressive variety of techniques have been applied to address this problem. Their breadth, and their limitations, serve to frame the difficulty of the task. Recalling the successes of intercalating dyes and covalently bound Fluorophore/Quencher, or FRET, pairs at monitoring the folding of tile systems one would expect similarly robust tools for larger DNA nanostructures. As discussed in chapter 3.1 that these techniques have only been rarely applied the folding of larger structures is telling both of the expense of FRET and of the complexity in scaffolded DNA nanostructures[11–13].

In part, the difficulty of examining scaffolded nanostructures stems from an inherent strength of DNA self-assembly; the rapid diffusion and motion of macromolecules that allows them to sample many states before locking into the correct one. The state of a nascent system can rapidly develop, and if taken out of equilibrium, the system may progress towards a metastable rather than equilibrium state. This is compounded by the relatively limited set of techniques which can image a 2 nm thick layer of carbon and nitrogen. Most typically characterization of DNA nanostructures is performed with AFM, TEM, or DNA-PAINT[14–16].The former two have not, to our knowledge, been applied to the folding process as they require specialized equipment and expertise.

One technique which has been used to observe folding is hot stage environmental AFM. This technique is conceptually simple; AFM is performed in a temperature and humidity

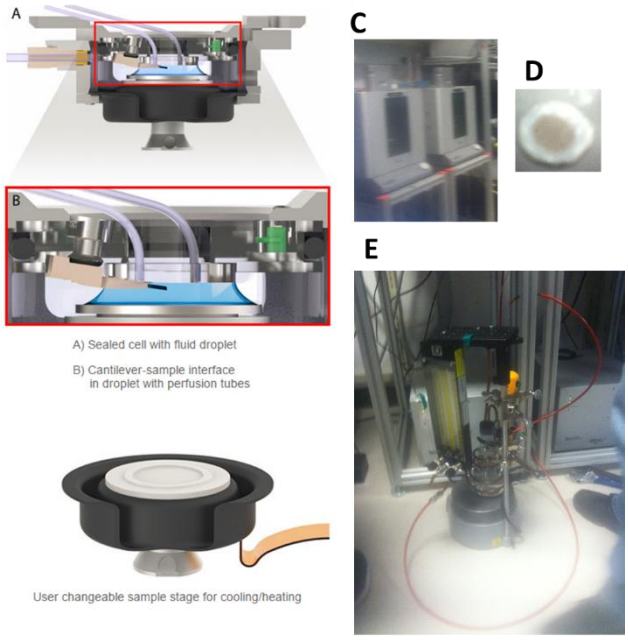


Figure 5.1: Hot Stage AFM. A-B, promotional material for Asylum hot stage environmental chamber, C-Cypher AFM, D-Sample puck, E-Custom humidity control

controlled chamber in liquid tapping mode. As shown in Figure 5.1 A-E, it often requires significant modification to an AFM; the annealing buffer must be isolated from the electronics, and it must be prevented from evaporating. Figure 5.1 A&B are promotional materials for the liquid environmental chamber sold by Asylum. Figure 5.1 C shows the Asylum system at the Center for Nanophase Materials Science, CNMS, at Oak Ridge National laboratory. The use of this system to monitor folding provides a window to the technical issues associate with hot stage AFM. As the sample chamber is hot and at high humidity, the mica disk is at risk for nucleating water bubbles between itself and the magnetic disk. Figure 5.1 D shows a mica disk bound by epoxy to a Teflon skirt, which is bound to the magnetic disk. Additionally, as environmental chambers are not always able to control humidity, custom made systems are often required, shown in Figure 5.1 E.

Temperature gradients leading from the stage can create imaging artifacts from both the tip and the stage motors. Additionally, the tight binding of the DNA nanostructure to the surface will significantly change the free energy of transition, and may altogether frustrate the folding process. After initial work with this technique, the Mingdong lab developed a set of clever experiments around these constraints[17,18]. They created nanostructures with intentional holes, then probed kinetics by measuring the time between staple introduction and hole filling as a function of temperature. To our knowledge, these are the only published studies utilizing hot stage AFM, in part due to the cost of modifying AFM equipment and in part due to the technical difficulties associated with imaging at high temperatures.

As discussed more thoroughly in section 4.2, the Turberfield lab created a dimer scaffold to probe folding by giving the staples a discrete set of choices during the folding process[11]. This work provides an excellent platform to examine folding preferences in general. However, as its insight into folding comes from a statistical distribution of chimeras, it does not readily reveal the folding pathway of a particular structure.

Another experiment which addressed the folding process is attempts to thermally quench DNA nanostructures. The Dietz lab, in the same publication as their work on RT-PCR fluorescence and rapid isothermal annealing, performed thermal quenches of scaffolded DNA nanostructures[9]. These quenches were not, to our knowledge, imaged. Rather, they were subjected to agarose gel electrophoresis to determine their annealing transition temperature as shown in Figure 5.2. Great care was taken in their protocols to minimize the time in which these systems were thawed and allowed to progress. This work was able to, qualitatively,

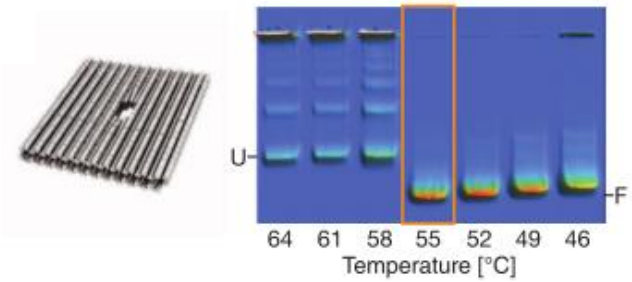


Figure 5.2: Gel image from Dietz isothermal annealing experiments[17,18]

show the development of a system from ssDNA to nanostructure, but as with fluorescence data, determination of folding pathways was beyond its intended scope.

It is plausible to create thermal quench protocols for imaging by purifying the separated bands after gel electrophoresis via freeze and squeeze. However, we chose not to pursue this due to inevitable complexity of such a protocol, and its propensity for human error. In particular, the propensity for these systems to progress once thawed would add significant time constraints on multiple steps through the procedure. It is worth noting that scaffolded nanostructures without staples, such as after freeze and squeeze, will be prone to denaturing after a relatively short period of time. This is in addition to the low yield, as low as 10%, and time consuming nature of freeze and squeeze. This complexity, risk, and reduced yield renders such a quench protocol less than feasible for general use.

The final technique we will discuss is the recent surface quench, described by the Estevez-Torres lab. This technique turns the disadvantages we discussed for hot stage AFM into an asset. The Estevez-Torres lab took low concentration, 1 nM, nanostructure samples of large volume ~1-1.5 mL and dipped mica disks into the samples mid-anneal. The samples were then washed and dried before imaging. Due to the strength of surface binding this effectively quenched the structures by freezing them to the mica surface. The Estevez-Torres lab chose to use the same model system, the Tall Rectangle origami or TR, as this work.

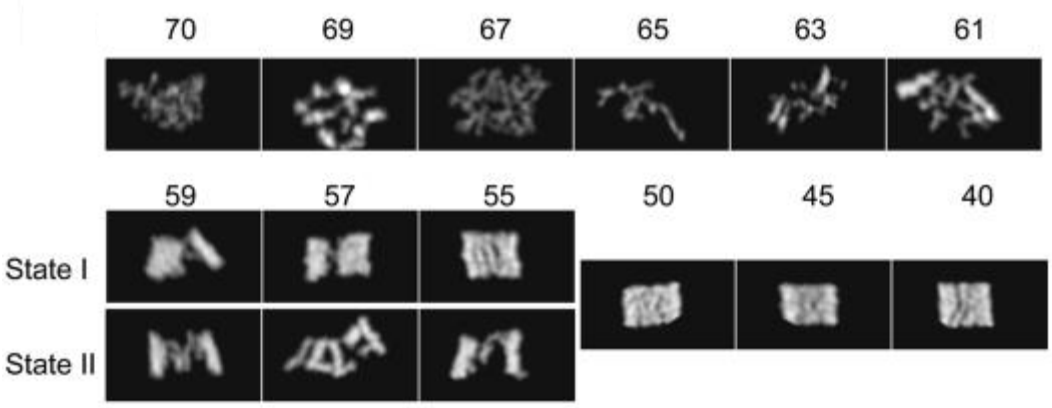


Figure 5.3: Reproduced AFM image from Estevez-Torres surface quenching study[32]

Figure 5.3 is a reproduction from the Estevez-Torres work. The images were not of high enough resolution to determine specific details of the folding. However, through the use of customized scripts, through the ImageJ[19] utility, they determined average structures as a function of temperature. They also used fluorescence spectroscopy data, located in their supplement, to find which annealing rates are at equilibrium and which were not. This aspect of their work is particularly valuable for further studies. It is quite exciting that the Estevez-Torres group directly observed the folding process of a DNA nanostructure without specialized equipment or complex protocols.

Unfortunately, the surface quench has several inherent limitations. The first being that only a narrow range of concentrations and staple excesses may be probed; too high of a nanostructure concentration or of staple excess would make imaging impossible. In order to insert an appropriately sized mica disk into the anneal tube, the anneal tube must contain ~1.5 mL of sample, or 500-1000x the standard volume. Additionally, the quenched sample will be sensitive to the motion of liquid across the surface of a mica disk; through surface tension, the drying process can cause damage to nanostructures. Finally, as the force driving quenching

is a surface interaction, this technique will be unreliable for 3D nanostructures and for early nucleation stages of Single Strand Tile based structures.

The diversity of attempts to monitor the entire folding process over the last 10 years reveal both the importance of this information and difficulty of obtaining it. Having described this previous work, section 5.2 will describe our proposed alternate technique which can be implemented in any lab with the equipment to perform basic DNA nanostructure experiments.

Section 5.3: Chemical Quench Requirements, and Candidates

In parallel to Esteves-Torres work, we propose to freeze DNA nanostructures changing the energetics of the system such that further annealing or melting is unfavorable. To do so, the structure must both be prevented from further base pairing and from melting/dehybridizing. As such, any modification must affect dsDNA and ssDNA differently. One way to do so is to remove one of the driving forces for base pairing by preventing hydrogen bonding in ssDNA, and to simultaneously to replace hydrogen bonding in dsDNA through a covalent crosslink. As this proposed protocol uses covalent reactions to quench the thermal state of the origami we call it the chemical quench.

For a chemical quench protocol to be valuable for the DNA nanotechnology community, it must be more than effective at probing folding conditions. It must be easily adopted by research groups without organic chemistry expertise. As such, it should use commercially available reagents have few steps. In turn, this requires a literature search for potential reagents from communities outside of DNA nanotechnology. Different research communities often implement different tools and evaluate work on different figure of merit which may not transfer to our needs. Searching the literature, identifying and testing potential reagents, and developing an easily adoptable protocol presents a time management and experimental design problem.

To streamline our protocol development, we chose to find and examine multiple reagents in parallel. To keep the initial search process focused, we created a list of 9 criteria, shown in Figure 5.4. Different reagents, evaluated by different research communities, will not have the same available information; removal of potential reagents which failed a criteria allowed for a rapid search and testing process. This was particularly useful when whole categories of reagents were removed from the search, such as heavy metal salts for generally deforming the dsDNA helix during crosslinking, criteria #4.

Evaluation Criterion for Chemical Quench Reagents:

1- *Must be commercially available and inexpensive*
2- *Must create adducts to render ssDNA unable to base pair*
3- *Must create crosslinks to render dsDNA unable to denature*
4- *Must not chemically degrade the DNA or alter its secondary structure*
   a. *phosphodiester cleavage/abasic sites*
   b. *excessive bending/opening of helix*
5- *Must react at temperatures between 25°C – 95°C*
6- *Must NOT be autocatalytic in water between 25°C – 95°C*
7- *Must be water soluble*
8- *Must not excessively denature DNA during chemical modification*
9- *Must be safe enough to use without <u>additional</u> personal protective equipment*

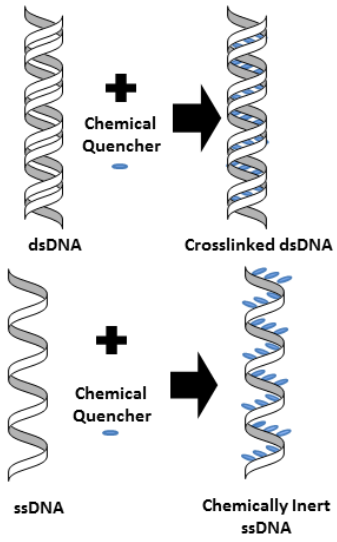


Figure 5.4: Evaluation Criterion and Schematic for Chemical Quench

In discussion of these criteria, it is worthwhile to recall the structure for Watson-Crick base pairing, as shown in Figure 5.5. The dotted lines in this figure show the hydrogen bonding which must be prevented for ssDNA or replaced with a crosslink for dsDNA. Regarding target moieties, it is worth recalling the distinction between primary, secondary, and tertiary moieties from introductory organic chemistry. These designations denote the number of carbon bonds to the moiety in question. A primary amine has one carbon bond, while a secondary amine has two carbon bonds, and a tertiary amine has three bonds with carbon atoms. This distinction is relevant as it greatly impacts reactivity; additional carbon bonds stabilize the

moiety in question, making the energetics of reaction less favorable. More aggressive reagents which may attack secondary and tertiary moieties are also more likely to attack phosphodiesters and cleave the DNA.

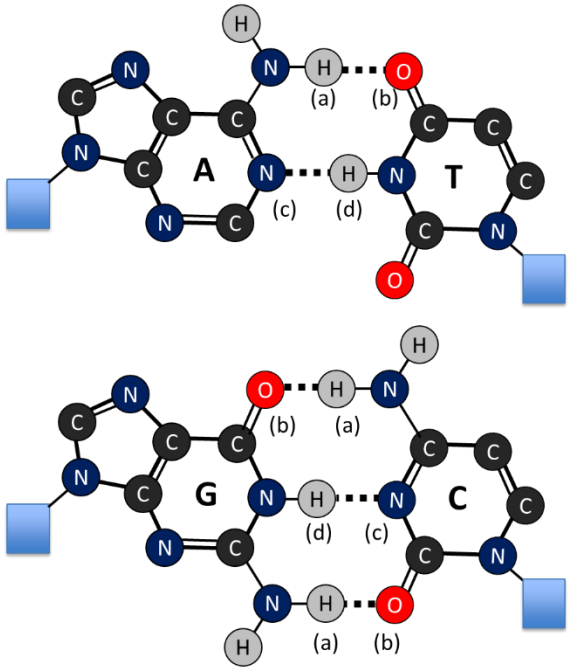


Figure 5.5: Watson-Crick base pairing, target moieties labeled

To interrupt base pairing, there are four moieties which one could target, shown in Figure 5.5. These targets are primary amines (a), the secondary carbonyls (b), tertiary amines (c), and secondary amines (d). The purines which are on left in Figure 5.5, have are multiple additional tertiary amines; attack of these sites will result in ring opening. Given the stability of secondary carbonyls and ring opening risk for tertiary amines, the literature search focused around conjugation chemistry for primary and secondary amines. Additionally, bifunctional linkers with high enough solubility to be used excess in were given priority; a bifunctional reagent in excess will crosslink dsDNA and cover ssDNA with enough bifunctional units to prevent crosslinks between ssDNA bases.

In discussing previous studies for interstrand DNA crosslinks and ssDNA adducts, it is important to consider the communities performing that research and how their priorities impact their work and figures of merit. In broad terms, three communities must be mentioned. The oncology community has studied into crosslinkers as chemotherapy agents, emphasizing those which can function in vivo and on reagents likely to trigger apoptosis, or cell death. The environmental safety community evaluates crosslinkers as safety hazards, again focusing on in vivo conditions, but with emphasis on negative outcomes rather than apoptosis. Finally, the microbiology community uses crosslinkers to fix living cells for study, with a focus on the reaction of the reagent with proteins. For all three communities, the exact physical structure of the link, the reactions with ssDNA, reactions in non-physiological buffers/temperatures, and water solubility are irrelevant to their figures of merit.

In short, all three communities from whom potential quenching reagents are drawn would not typically publish information useful to criterion #3-7. Additionally, all three communities tend to consider all damage to DNA in the same context, in terms of its ability to affect the cell. For the proposed chemical quench, all DNA damage is not created equal. Some DNA damage is necessary to freeze the structure, but some is undesirable as it could warp the nanostructures.

After initial evaluation 5 promising reagents were picked, two from the fixing community, one from the chemotherapy community, one from the environmental hazard community, and one crosslinker used on DNA nanostructures in the past[20]. These were glutaraldehyde, formaldehyde, streptozotocin, epichlorohydrin and psoralen, shown in Figure 5.[21–26].

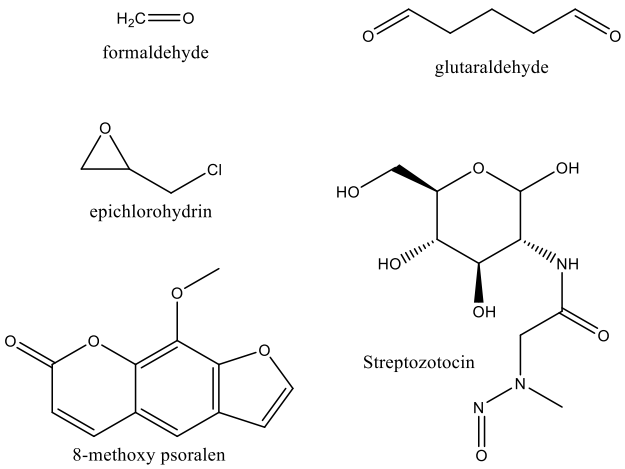


Figure 5.6: Potential quenching reagents

Psoralen, more specifically 8-methoxy psoralen or 8MOP, intercalates into DNA and creates AT/TA crosslinks after UV exposure[20,27]. These crosslinks have been demonstrated to not destroy DNA nanostructures. It is unknown if 8MOP adducts prevent ssDNA hybridization. Epichlorohydrin, glutaraldehyde, formaldehyde, and streptozotocin are all known for causing DNA crosslinks and adducts[21,24,25,28–30].

In this section we have discussed our initial criteria for chemical quenching reagents, and outlined our choices for promising reagents. In section 5.4, we will address our procedures and initial screening for narrowing the number of potential quenching reagents.

Section 5.4: Initial Screening and Experimental procedure

Of the reagents examined, formaldehyde, glutaraldehyde and streptozotocin are all water soluble, while epichlorohydrin (ECH) and 8-methoxy Psoralen (8MOP) are only sparingly water soluble. Both 8MOP and ECH were solubilized with water by dilution into dimethyl sulfoxide (DMSO) before addition to annealing buffer. This presents some safety concern as DMSO is able to rapidly diffuse through skin, and will readily carry dangerous molecules with it; this ability is significantly reduced when the DMSO is mixed with water.

These concerns were addressed by only preparing 10-75 µL of DMSO stock at a time and by double gloving. Any DMSO stock not in use were put into chemical waste immediately.

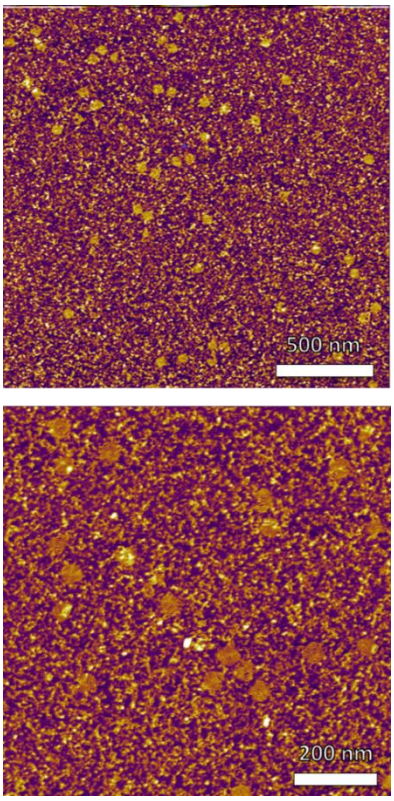


Figure 5.7: Initial 8-Methoxy Psoralen addition and crosslinking at 90°C

8MOP has been used to crosslink DNA nanostructures in the literature, and its UV activation requires additional implementation. It was tested first, for these reasons. Three samples of circle miniM13 origami were made at 15 nM in 1x TAE Mg buffer, and tested for criterion #2 – rendering ssDNA unable to anneal. All three samples and the TAE Mg/DMSO/8MOP stock were brought to 90°C. Then excesses of 8MOP were added, of values 0x, 500x, and the solubility limit, of 8MOP. All samples were then irradiated with UV-A light for 1 hr. The samples were then imaged, and all showed formation of the circle nanostructures as shown in Figure 5.7. As it fails criteria #4, 8MOP was ruled out as a quenching reagent.

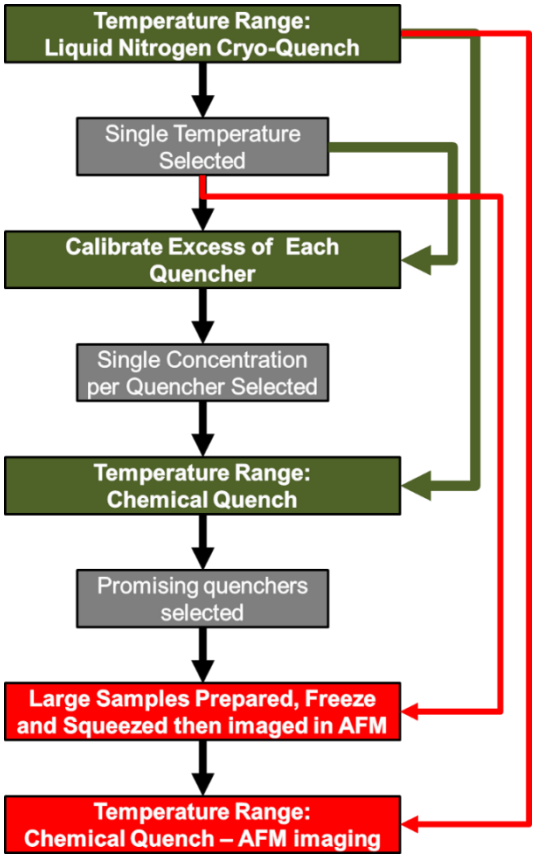


Figure 5.8: Protocol refinement flowchart

Following this, screening was performed as described in the flow chart, Figure 5.8 . We had to choose a systematic screening process due to the significant number of variables involved in making an effective protocol. Following this path allowed us to progress logically, with clear data for comparison from previous steps. Another feature of this experimental flow path is that it allowed us to minimize AFM characterization during early stages when large sets of variables were being compared. In this chart green indicates characterization via gel electrophoresis, grey indicates decisions to narrow a single variable, and red indicates characterization via AFM.

Liquid nitrogen quenches were performed using protocols published by the Dietz lab[9]. The PCR unit was paused then, as quickly as possible, the tubes were transferred into liquid

nitrogen and the unit un-paused. After spending at least one minute in the liquid nitrogen, the quenched tubes were transferred to the freezer.

After all the samples were quenched, a 2% agarose gel was prepared in TAE Mg buffer. The samples were then thawed, mixed with loading dye, and loaded into the gel as quickly as possible. The gels were then run at 4°C for 6 hr at 2.5 V/cm then overnight, approximately 8.3 hr at 1.25 V/cm. Staining was performed with Ethidium Bromide, with the resulting image shown in Figure 5.9.

In Figure 5.9 secondary bands can be seen bordered by the red dotted lines. These bands are a clear parallel to those seen by the Dietz lab, particularly in that they disappear when the lower band transitions to the level of the fully formed origami. Also in Figure 5.9, an

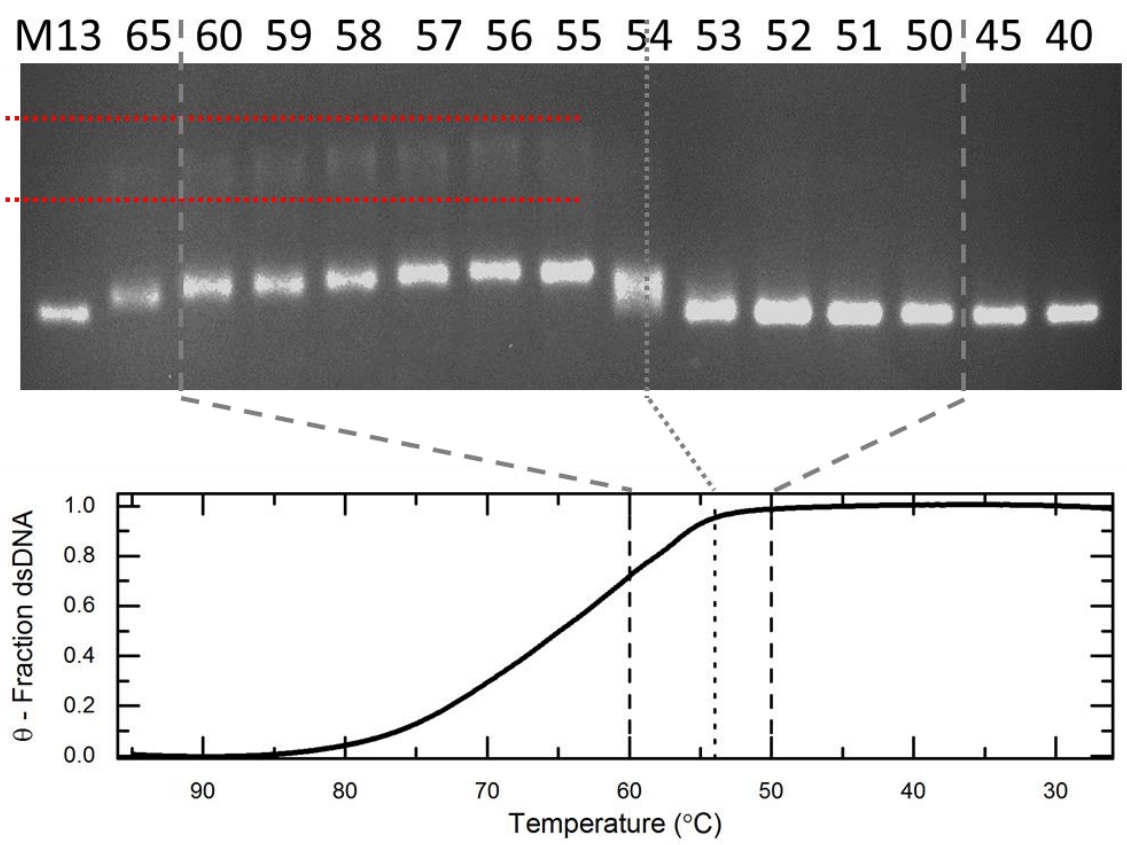


Figure 5.9: Liquid nitrogen quench and anneal curve for Tall Rectangle (TR)

anneal curve with the temperature scale starting at 95°C and going down to 25°C is shown. The transition point, ~54°C maps almost perfectly onto the leveling of the anneal curve.

From Figure 5.9, 56°C was selected as the reference temperature for calibration of quencher reagents as per Figure 5.8. 56°C was chosen because it had secondary bands and was close to the transition. To calibrate the excess of quencher, the overall number of nucleobases in a 25 µL anneal was calculated, including excess ssDNA staples. A quencher's number of times excess was in reference to that number. A set of stock solutions for each potential quencher was adjusted so that addition of 7.5 µL would provide 1x, 20x, 100x, 500x, 1000x, and 3000x of each stock.

Deliberate care was taken in following the same protocol for addition of each quencher. At the desired temperature, the PCR unit was paused, and as quickly as possible, the micropipetter was used to draw and expel 7.5 µL of quencher solution 25 times to equilibrate the temperature of the pipette tip. The remaining volume was then fully expelled, and a fresh 7.5 µL of quencher was drawn into the pipette tip. The sample tube was then opened, and the quencher added, while moving the tip in a circular motion, simultaneously drawing and expelling 25 times. In transition from the quencher tube to the sample tube the pipette tip was visibly inspected to check that the volume was appropriate; this was necessitated by the occasional mis-draws caused by micropipetting a hot, evaporating solution.

The PCR unit was left paused for 5 minutes, and then unpaused to continue annealing. The anneal rate was set to 0.21°C/min, or approximately 4.6 hrs from 80°C to 25°C. The resulting samples were run in a 2% agarose gel in 1x TAE Mg buffer, for 6 hr at 2.5 V/cm all at 4°C. Ethidium bromide was used for staining.

As shown in Figure 5.10, the quench had varying degrees of impact. The formaldehyde and glutaraldehyde varied most from pure TR. Streptozotocin appeared to have little to no

effect on the samples, this is likely because of autocatalysis at high temperatures; Streptozotocin is known to degrade readily, so this was not entirely unexpected[31]. Epichlorohydrin showed some promise, as its bands were slightly displaced from the TR, and there was a faint secondary band. The missing lane for 20x epichlorohydrin is due to poor loading of the gel, this set was not rerun, as the missing band does not affect the information imparted by the gel.

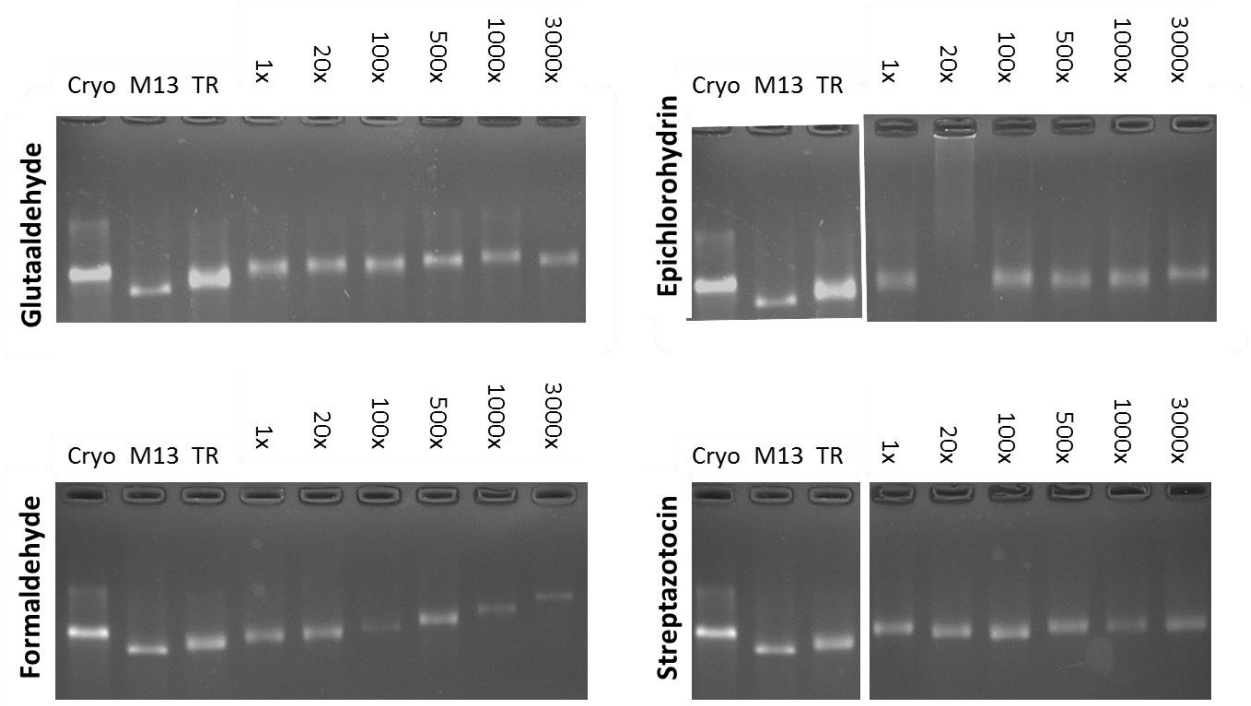


Figure 5.10: Excess quencher added to TR at 56°C for all samples

From these results, streptozotocin was removed from further consideration as a chemical quenching reagent. As there was no clear degradation of the DNA at any concentration for glutaraldehyde or formaldehyde, 500x was chosen as a reasonable excess.

The chemical quench procedure was then repeated over the same temperature range as the liquid nitrogen range for 500x glutaraldehyde, formaldehyde and epichlorohydrin. After quenching then allowing the anneals to finish, the samples were run in a 2% agarose gel

under the same conditions as the liquid nitrogen quench system. These gels are shown in Figure 5.11.

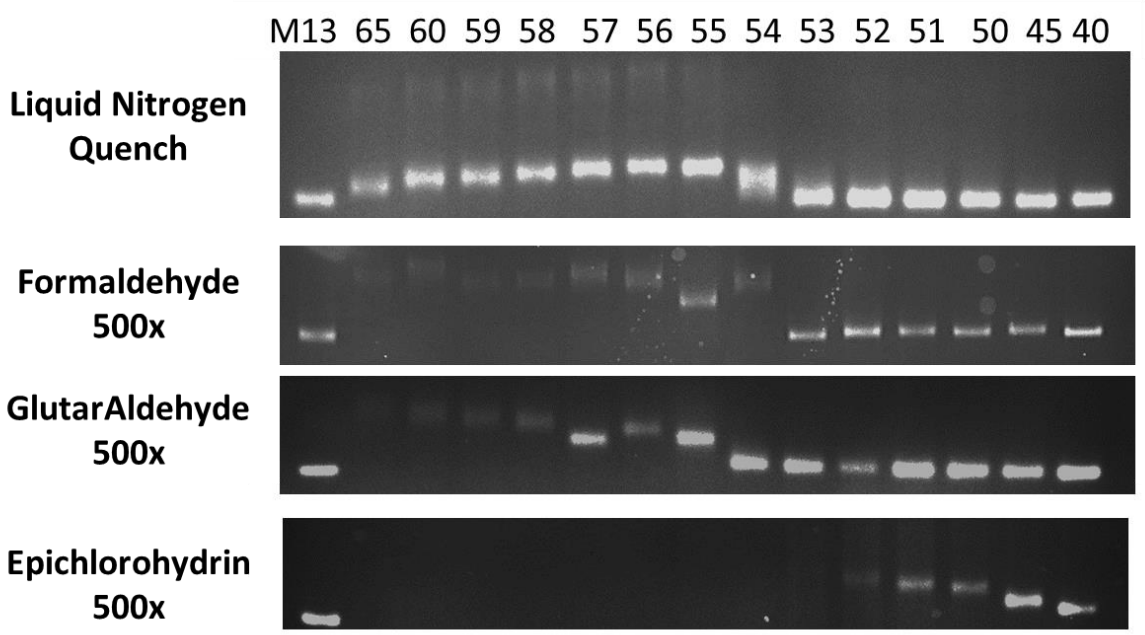


Figure 5.11: Liquid nitrogen and chemical quenches across temperature range

As shown in Figure 5.11, both the formaldehyde and glutaraldehyde anneals have a transition between 55-54°C, in parallel to the liquid nitrogen quench. It is of note that the epichlorohydrin appears to completely degrade the DNA at temperatures well below those shown in Figure 5.10. Some error must exist in the epichlorohydrin samples between Figure 5.10 and Figure 5.11. Regardless of which experiment was in error, the energy stored in the epoxide group was already a concern for cleaving phosphodiester bonds along the DNA backbone. Given the clear transition in the formaldehyde and glutaraldehyde that parallels the transition for liquid nitrogen, we chose to remove the epichlorohydrin from consideration as a quenching agent. This had the benefit of eliminating the safety hazard of DMSO.

The choice to remove epichlorohydrin from consideration left only two potential quench conditions for a preliminary protocol, namely formaldehyde and glutaraldehyde at 500x

excess. This is a small enough sample set to make AFM imaging feasible. We created three large, 50 µL anneals. One for a liquid nitrogen quench, one for formaldehyde, and one for glutaraldehyde; all were quenched at 56°C. These samples were then run in a 2% agarose gel as before, then freeze and squeeze protocols were used to purify the samples from the gel. Given the low yield of freeze and squeeze, often as low as 10%, large volume anneals was necessary. The resulting images are shown in Figure 5.12.

A relevant observation from the liquid nitrogen quench is the structure from the upper, or secondary, band. The structures shown in this band clearly indicate intertwined scaffolds, which are likely an artifact of either system development after thawing, or the delay in moving tubes from the anneal into liquid nitrogen. Because these secondary bands appear in other studies, it is likely that these artifacts are one of the limitations of the nitrogen quench[9]. The primary band of the liquid nitrogen quench parallels mid-formation stages of the Estevez-Torres work, albeit at different temperatures[32]. The temperature difference is likely due to the

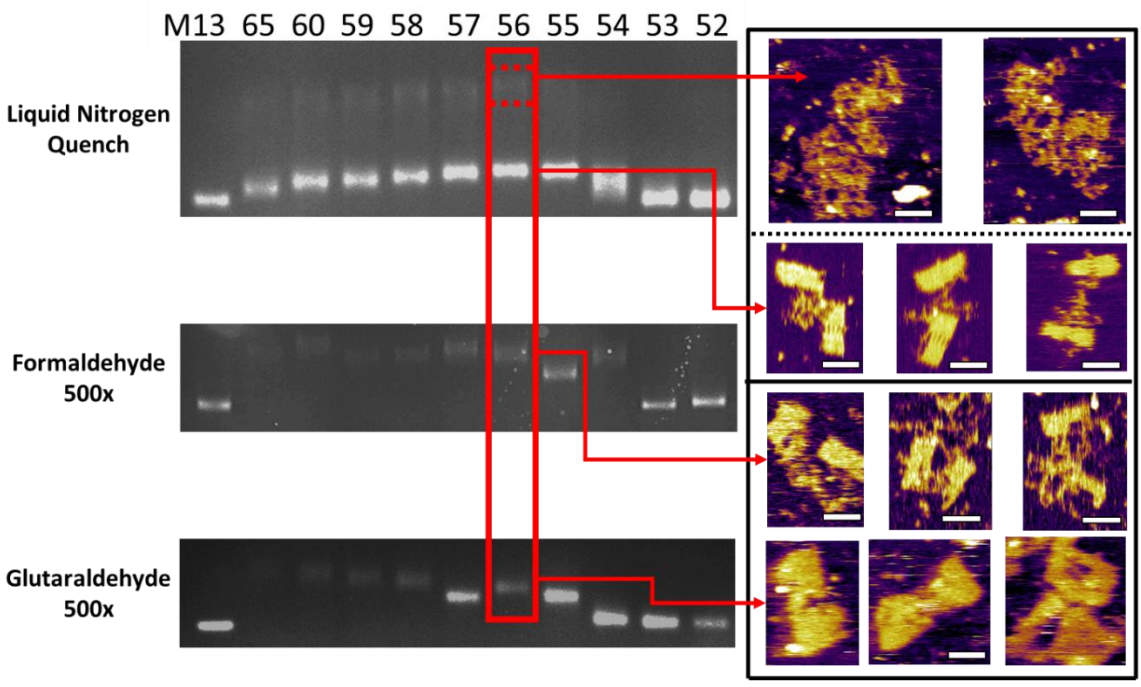


Figure 5.12: Freeze and Squeeze of liquid nitrogen, formaldehyde, and glutaraldehyde quenches at 56°C

much higher concentration in these samples, which should increase the transition temperature for annealing. The extracted formaldehyde band appears quite similar to the liquid nitrogen primary band, with slightly less structure formation, while the glutaraldehyde band is clearly further progressed.

The error in thawed liquid nitrogen anneals is towards more developed structures, while a chemical quencher is likely to denature a system on quenching. Between these, it is impossible to tell whether the formaldehyde or liquid nitrogen quenches are "true" representations of the TR system at 56°C. However, it is reasonable to use these techniques to probe the folding order of the nanostructure. Given the advanced progression of the glutaraldehyde sample, it is likely that more quencher excess would be necessary for that reagent. However, as the formaldehyde quench seems sufficient at 500x we chose to use formaldehyde for our preliminary protocol.

In this section we have iterated through a series of tests to define our preliminary protocol for chemical quenching. In section 5.5, we will apply this preliminary protocol to the formation of the TR nanostructures.

Section 5.5: Application of the Preliminary Chemical Quench

The protocol described in section 5.4 was used with 500x excess formaldehyde and the quenched samples were imaged without freeze and squeeze purification. Figure 5.13 shows the routing map, circle map, and GC content map of the TR.

Figure 5.14 shows the AFM results for these samples. From these structures, the formation of the TR is clearly evident. The loop of excess ssDNA is visible in the 57°C sample, often this is referred to as the "goatee." This goatee allows us to infer that the bottom half of the TR forms first. From the GC map, Figure 5.14 far right, it is clear that there are regions at both the top and bottom of the TR which have high GC content. The placement of this content

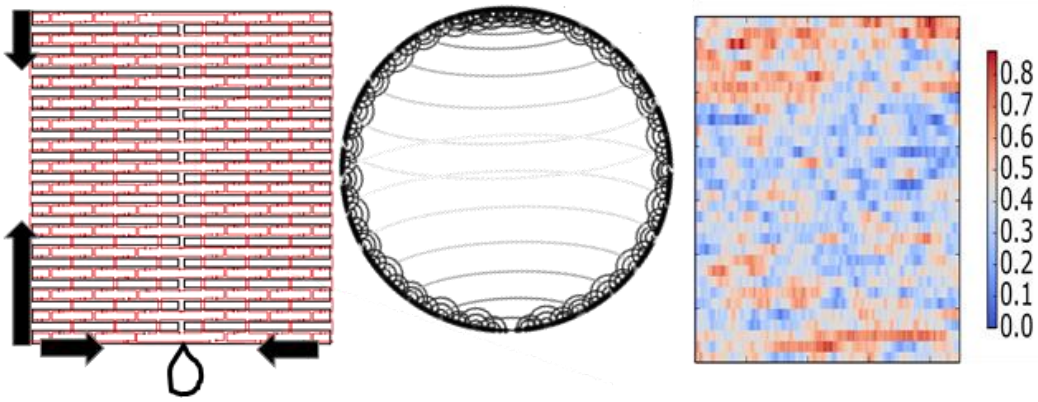


Figure 5.13: Routing map, circle map, and GC content map of Tall Rectangle. In GC map, red indicates high GC content, blue indicates high AT content

provides useful insight; the top half of the TR has generally more GC content but it is spread out, while the bottom half of TR has a line of high GC content which bridges the seam. As the bottom half clearly forms first, we can reasonably argue that GC content itself is not the only factor in nucleation, the location of GC content is vital.

Figure 5.13., shows the routing map, circle map, and GC content map for the TR system. In Figure 5.13, red equates to high GC content, while blue equates to high AT content. Barring other factors, one would anticipate the red regions to anneal first. The arrows on the routing map of the TR indicate the annealing pathway preferences of the TR structure. From Figure 5.14 58°C and 57°C, we can infer that the staple binding pads bind independently at high temperature, followed by closure of staples staring with those which bind relatively close sections of the scaffold. This makes the nanostructure form from the top, bottom, and sides inwards to the center.

In this section, we have applied our preliminary protocol to the TR system, and placed our results in the context of the TR circle map and GC content map. In section 5.6 we will discuss our steps to finalize the chemical quench protocol.

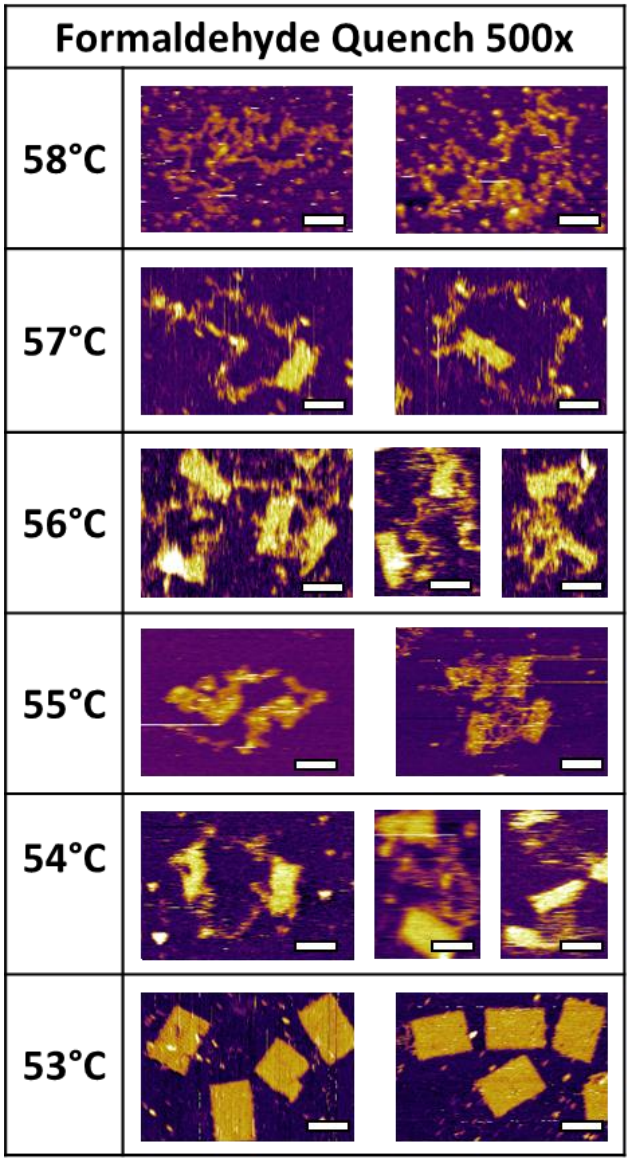


Figure 5.14: AFM images of formaldehyde quenched Tall Rectangle

Section 5.6: Protocol finalization

The protocol described in section 5.4 was implemented on several miniM13 origami structures with limited success. Given the limited number of attempts, it is possible that this

error comes from the inherent variability of AFM imaging, the chemical quench could have a limited viable temperature range, or some difference between the structures makes the quench destructive to the DNA.

While further tests are required, this provides an opportunity to discuss the mechanism of crosslinking, and finalization of the quenching protocol. While there is still debate on the

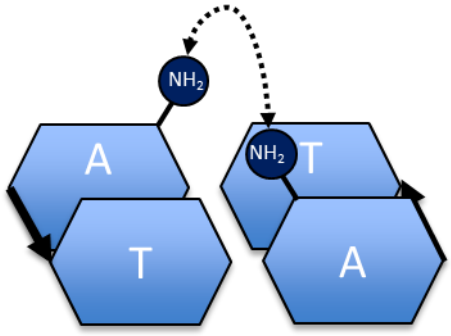


Figure 5.15: Schematic of most common interstrand formaldehyde crosslink

exact reactions between formaldehyde and DNA, there is general consensus that the most common crosslinks occur in AT/TA sequence pairs, shown in Figure 5.15. However, it has been indicated that at least some of the GC bases are also crosslinked, although with less regularity than AT/TA pairs. This is reasonable, as G and C bases both have primary amines in similar positions to A bases.

As the number of primary amines on each nucleobase is different, the required quencher excess could need to be changed between scaffolds. Towards this end we calculated the number of primary amines in both scaffolds, as well as the number of primary amines in a full anneal, including staples. The distinction between the number of primary amines in an anneal and in the scaffold is vital. The A and T bases do not have identical numbers of primary amines, and the 10x excess staples in a standard anneal should consume a significant proportion of the total formaldehyde. As a result, not only is the AT/GC ratio important, but also the A/T ratio in the scaffold. Table 5. shows our estimates of the number

of adenine bases and primary amines for the minM13 and M13MP18 scaffolds. This indicates that, barring some unseen but vital difference between the miniM13 and M13 scaffolds, the quench protocol for the two scaffolds should be identical.

Table 5.1:Comparative scaffold GC content, number of adenine bases, and number of primary amines M13 and miniM13

	%GC	Length	'A' bases	Primary Amines Scaffold	Primary Amines Anneal
miniM13	49%	2,404	594	1,773	19,873
M13MP18	42%	7,249	1,766	4,830	59,660
	Ratio	0.33	0.34	0.37	0.33

As a next step, initial tests to calibrate the excess of formaldehyde were created. Quenches were performed with 1°C increments between 55-50°C at quencher excesses of 200x, 400x, 800x, and 1,600x. We show these, with the 500x quench from Figure 5.12, in Figure 5.16, as this sample was quenched in steps, it did not have a pure TR band for comparison. Figure 5.16 shows that the 500x used in the preliminary protocol was reasonable.

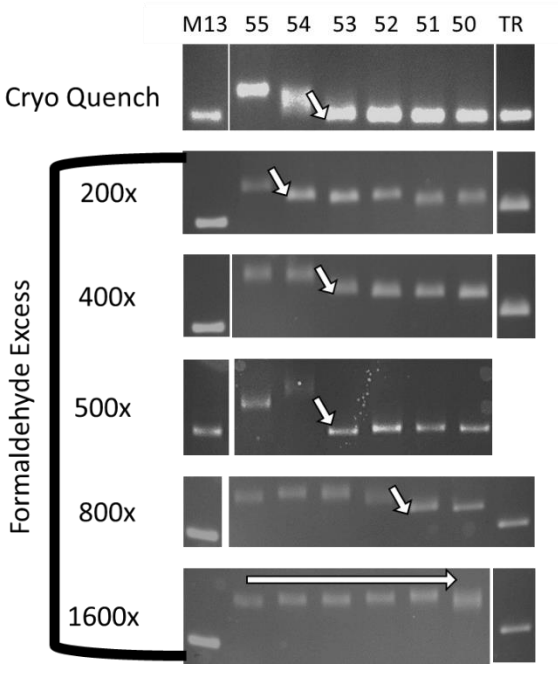


Figure 5.16: Initial calibration experiment for excess formaldehyde

The final possibility which must be addressed for protocol finalization is that the reaction rates for DNA and formaldehyde at different temperatures could differ significantly. The quencher excess may need to vary to compensate. In the future work section, we will discuss our proposed experiments to determine whether excess variation is necessary and how to calibrate it.

In this section we have discussed the steps we have taken towards making the chemical quench protocol robust enough to be applied in any lab with an annealer, formaldehyde and micro-pippetters. In section 5.6, we will discuss the test system which we built to probe, with chemical quenching and fluorescence spectroscopy, the link between nanostructure design and its folding.

Section 5.7: Engineering Folding Pathways

The ability to effectively quench nanostructures is exciting and will open the door to important new studies relating yield and folding; to best demonstrate its capability, we created a model system to emphasize what we anticipate will be differences in folding for structurally identical shapes.

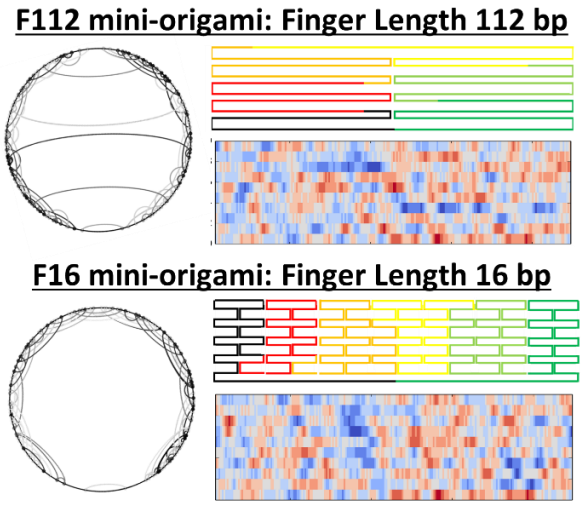


Figure 5.17: Test system for scaffold routing. Scaffold routing maps, circle maps, and GC content maps. Red indicates high GC content

To create such a model system, we again turned to the impressive prototyping capabilities of the miniM13 scaffold. Two scaffolded structures were designed, the Finger 112 and Finger 16 structures, or F112 & F16 respectively. Finger refers to the number of bases in each scaffold "finger" before it folds back on itself.

The F112 is identical to the TR in all dimensions except its height. The F16 had an identical footprint to the F112, but a radically different routing pattern. This was achieved by replacing the staple crossovers with scaffold crossovers. In doing so, the distribution of distances along the scaffold, bridged by the staples, became much more gradual and centered at a shorter distance. These distinctions are shown in Figure 5.17 with which also shows GC content maps which indicate little significant change in the clustering of high or low GC regions.

Figure 5.18 shows AFM images of the F16 and F112 structures with their anneal curves. It is worth noting that for the dotted anneal curves, the edge staples were left out of the structure. This ensured that both structures had an identical number of total crossovers, which was not the case if the edge staples were included. For both anneal curves with and without sides, the F16 annealed at a higher temperature than the F112. This provides evidence of the role of design in thermal properties, and will provide a valuable test case for the chemical quench.

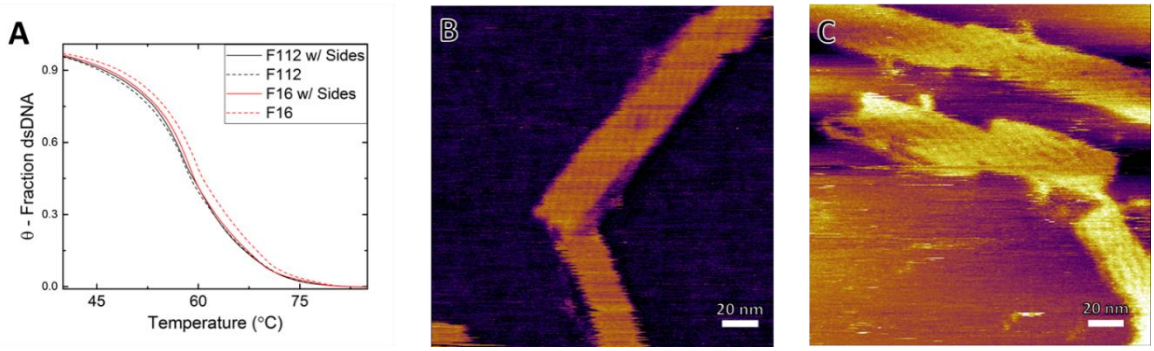


Figure 5.18: A-anneal curves for F112 & F16, with and without side staples, and AFM images for B-F16 nanostructures, and C-F112 nanostructures

In this section we have discussed the model F112 and F16 DNA nanostructures, and their design which acts to highlight how scaffold routing can affect thermal properties. In section 5.7, we will discuss our conclusions from this chapter, and our future work on this topic.

Section 5.7: Conclusions and Future Work

In this chapter we have introduced the concept of quenching DNA nanostructures, section 5.1; discussed the numerous attempts to characterize folding pathways, section 5.2; introduced the chemical quench and set parameters in our search for an appropriate protocol, section 5.3; described our screening experiments which led to a proposed protocol, section 5.4; applied that protocol to the standard model Tall Rectangle DNA nanostructure, section 5.5; discussed the final steps required to cement the protocol, section 5.6; and introduced a model system which we will use to demonstrate the value of the chemical quench protocol, section 5.7.

As a next step, we will gather the information regarding variation of the formaldehyde excess as a function of temperature. We will order a simple pair of complimentary ssDNA strands, and across the temperature range of 95°C-25°C we will repeat two tests. One test will determine if a given excess of formaldehyde is sufficient to prevent ssDNA from hybridizing with its compliment and the other test will determine if the excess is sufficient to crosslink dsDNA. These tests will be analyzed with gel electrophoresis, which will also show evidence of phosphodiester cleavage should it occur.

These two tests will be initiated at the temperature extremes, 95°C and 25°C with the tests progressing inwards to 55°C. By using system of short ssDNA and dsDNA, we can use Poly Acrylaminde Gel Electrophoresis (PAGE), which will allow higher voltages, shorter run times, and cleaner images.

After finalization of the quench procedure, the quench will be performed first on the F112 and F16 nanostructures. Once this has been completed, the multishape DNA nanostructure anneals from Chapter 4 will be quenched, allowing a complete test of the nucleation/growth mechanism.

References

1. Gállego, I., Grover, M. a. & Hud, N. V. Folding and Imaging of DNA Nanostructures in Anhydrous and Hydrated Deep-Eutectic Solvents. *Angew. Chemie Int. Ed.* n/a–n/a (2015). doi:10.1002/anie.201412354

2. Lee, J. *et al.* The restoration of DNA structures by the dry–wet method. *Soft Matter* **8,** 619 (2012).

3. Kim, H., Surwade, S. P., Powell, A., Donnell, C. O. & Liu, H. Stability of DNA Origami Nanostructure under Diverse Chemical Environments. (2014).

4. Campos, R., Zhang, S., Majikes, J. M., Ferraz, L. C. C. & Labean, T. H. Electronically addressable nanomechanical switching of i-motif DNA origami assembled on basal plane HOPG. *Chem. Commun.* **51,** 14111–14114 (2015).

5. Shin, J.-S. & Pierce, N. a. A synthetic DNA walker for molecular transport. *J. Am. Chem. Soc.* **126,** 10834–5 (2004).

6. Simmel, F. C. Processive motion of bipedal DNA walkers. *Chemphyschem* **10,** 2593–7 (2009).

7. Ouldridge, T. E. *et al.* Optimizing DNA Nanotechnology through Coarse-Grained Modeling :Walker, A Two-footed D N A. *ACS Nano* **7,** 2479–2490 (2013).

8. Wang, Z.-G., Elbaz, J. & Willner, I. DNA machines: bipedal walker and stepper. *Nano Lett.* **11,** 304–9 (2011).

9. Sobczak, J.-P. J., Martin, T. G., Gerling, T. & Dietz, H. Rapid folding of DNA into nanoscale shapes at constant temperature. *Science* **338,** 1458–61 (2012).

10. Zhang, Z., Song, J., Besenbacher, F., Dong, M. & Gothelf, K. V. Self-assembly of DNA origami and single-stranded tile structures at room temperature. *Angew. Chemie - Int. Ed.* **52,** 9219–9223 (2013).

11. Dunn, K. E. *et al.* Guiding the folding pathway of DNA origami. *Nature* (2015). doi:10.1038/nature14860

12. Nangreave, J., Yan, H. & Liu, Y. Studies of thermal stability of multivalent DNA hybridization in a nanostructured system. *Biophys. J.* **97,** 563–571 (2009).

13. Wei, X., Nangreave, J., Jiang, S., Yan, H. & Liu, Y. Mapping the thermal behavior of DNA origami nanostructures. *J. Am. Chem. Soc.* **135,** 6165–6176 (2013).

14. Jungmann, R. *et al.* Multiplexed 3D cellular super-resolution imaging with DNA-PAINT

and Exchange-PAINT. *Nat. Methods* **11,** 313–8 (2014).

15. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).

16. Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325,** 725–730 (2009).

17. Song, J. *et al.* Direct visualization of transient thermal response of a DNA origami. *J. Am. Chem. Soc.* **134,** 9844–7 (2012).

18. Song, J. *et al.* Isothermal hybridization kinetics of DNA assembly of two-dimensional DNA origami. *Small* **9,** 2954–2959 (2013).

19. Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with imageJ. *Biophotonics International* **11,** 36–41 (2004).

20. Rajendran, A., Endo, M., Katsuda, Y., Hidaka, K. & Sugiyama, H. Photo-Cross-Linking-Assisted Thermal Stability of DNA Origami Structures and Its Application for Higher-Temperature Self-Assembly TL - 133. *J. Am. Chem. Soc.* **133 VN - ,** 14488–14491 (2011).

21. Romano, K. P., Newman, A. G., Zahran, R. W. & Millard, J. T. DNA interstrand cross-linking by epichlorohydrin. *Chem. Res. Toxicol.* **20,** 832–838 (2007).

22. Olsen, R., Backman, J., Molander, P., Klika, K. D. & Kronberg, L. Identification of Adducts Formed in the Reactions of 2′-Deoxyguanosine and Calf Thymus DNA with Glutaraldehyde. *European J. Org. Chem.* **2007,** 4011–4018 (2007).

23. Taatjes, D. J. & Walker, J. M. *Cell Imaging Techniques*. **319,** (2006).

24. Cheng, G. *et al.* Reactions of Formaldehyde Plus Acetaldehyde with Deoxyguanosine and DNA : Formation of Cyclic Deoxyguanosine Adducts and Formaldehyde Cross-Links. 145–152 (2003).

25. Cheng, G., Wang, M., Upadhyaya, P., Villalta, P. W. & Hecht, S. S. Formation of Formaldehyde Adducts in the Reactions of DNA and Deoxyribonucleosides with r -Acetates of N -Nitrosodimethylamine (NDMA). 746–751 (2008).

26. Voulgaridou, G.-P., Anestopoulos, I., Franco, R., Panayiotidis, M. I. & Pappa, A. DNA damage induced by endogenous aldehydes: current state of knowledge. *Mutat. Res.* **711,** 13–27 (2011).

27. Shi, Y. *et al.* in *Bioorganic photochemistry: photochemistry and the nucleic acids* (ed. H., M.) (Wiley, 1990).

28. Huang, H. & Hopkins, P. B. DNA Interstrand Cross-Linking by Formaldehyde:

Nucleotide Sequence Preference and Covalent Structure of the Predominant Cross-Link Formed in Synthetic Oligonucleotides. 9402–9408 (1993).

29. Hermanson, G. T. Bioconjugation techniques. *Acad. Press* **10,** 0123705010 (2008).

30. Sund, P. & Kronberg, L. Reaction of epichlorohydrin with adenosine, 2???-deoxyadenosine and calf thymus DNA: Identification of adducts. *Bioorg. Chem.* **34,** 115–130 (2006).

31. Lee, J. Y., Kim, M. J., Moon, C. K. & Chung, J. H. Degradation products of streptozotocin do not induce hyperglycemia in rats. *Biochem. Pharmacol.* **46,** 2111–2113 (1993).

32. Estevez-torres, A. Observing and Controlling the Folding Pathway of DNA Origami at the Nanoscale. (2016). doi:10.1021/acsnano.5b05972

Chapter 6: Conclusions and Future Work

Section 6.1: Summary of Conclusions

This work was motivated by the need for better understanding of, and better protocols for examining, DNA nanostructure folding. Structure folding mediates production yield and, we have shown, is dependent on design. As DNA nanotechnology moves from the lab bench to industry these techniques and the knowledge they create will inform both design choices and yield optimization.

We developed four new techniques, the miniM13 workbench scaffold for low cost rapid prototyping; implementation of baseline correction, including photobleaching, for fluorescence spectroscopy annealing and melting data; competitive multishape annealing for determining the comparative favorability of differing shapes; and chemical quenching to allow direct examination of folding processes.

These techniques, and their applications in model systems led to the following insights.

- miniM13 workbench scaffold
    - Upfront costs and Material costs can and should be treated separately for nanostructure systems
    - Reductions in upfront costs can significantly increase the number of structures available for experimental examination
- Fluorescence spectroscopy
    - It is possible to use fluorescence spectroscopy data to determine the hysteresis area of DNA nanostructures

    - The hysteresis of the Tall Rectangle model system, and thus its folding/unfolding cooperativity, appear to have a local minimum as a function of reagent excess
    - Photobleaching can and does occur in annealing and melting of DNA nanostructure systems and can be corrected for
- Multishape competitive anneals
    - Structures formed in multishape competitive anneals can be rationally understood and controlled, at least in part, by design
    - Strand displacement pushes competitive systems to stable conformational states
    - Phase diagrams can be predicted for competitive systems using anneal data
    - Nucleation and growth is a reasonable model for nanostructure formation
- Chemical quenching
    - Chemical quenching of DNA nanostructures is a reasonable method for examining folding pathways
    - The Tall Rectangle model system shows that folding favors staples with small reductions in conformational entropy
        - These positions can be interpreted as nucleation points
    - Design, specifically routing, is shown to definitively affect structure folding in the F16/F112 system
    - Spreading the contribution of conformational entropy reductions more evenly across all of the staples results in a structure which anneals at a higher temperature, as evidenced by the F16/F112 system

Collectively the tools we have developed provide not only insight into DNA nanostructure folding, but also provide synergy with each other. By gaining ensemble level data from fluorescence spectroscopy we are able to examine broad trends, such as hysteresis, that were previously difficult to observe. These broad trends then inform our understanding of molecular folding systems in competition, and support the

hypothesis that nucleation and growth is a good model for DNA nanostructure formation. Finally, both of these help to provide context for the folding snapshots provided by the chemical quench. In turn, the chemical quench will provide concrete information regarding the progression of competitive systems and will provide details unattainable by ensemble anneal/melt curves.

Application of these tools in concert, in our lab and others, should provide exciting new information regarding folding and design of DNA nanostructures moving forward.

Section 6.2: Future Work

Given the insights and protocols described above, we propose the following future work as both initial next steps and longer term objectives.

- Immediate Next Steps:
    - Finalize data correction for photobleaching – Run a 96 well plate of pH, photobleaching and staple excess samples and fit photobleaching function.
    - Perform AFM of 3-6 test positions for competitive anneals
    - Finalize chemical quench – determine if temperature dependence of reaction rate merits adjustment of formaldehyde excess as a function of temperature
    - Chemically quench model F112 and F16 samples
    - Chemically quench competitive anneal system
- Long term objectives:
    - Create general guidelines for the relation of scaffold routing and cooperativity/annealing temperature
    - Create general guidelines for the relation of staple motifs and cooperativity/annealing temperature
    - Examine assembly yield as a function of system cooperativity and scaffold routing
    - Test “super” staples in competitive anneals for forcing systems into different folding pathways
    - Create planned chimera shapes using GC content mapping to overlay high GC content onto staples which significantly reduce conformational entropy

Addressing both short and long term recommendations for future work will improve both the thermodynamics toolkit for DNA nanostructure, our understanding of design/property relationships, and, over the long term, the yield of DNA nanostructures.

APPENDICES

Appendix 1: Data analysis script for Quantstudio multicomponent data

(require Python and Scipy- anaconda package used)

```
#organize by wellname
import sys; import copy; import numpy as np; import scipy.stats as stats; import scipy as sy
import os; from operator import itemgetter; import scipy.optimize as sp; import math; import
modelFunctions
hideMe=1 # helps to organize the script
fname="anneal06CS.txt"
verbose=0 # verbose =0 supresses extra printing to the command prompt
print("This script will run on the multicomponent and sample setup data from quantStudio
Equipment")
print("Please copy your sample setup and multicomponent text files, then remove all lines
before the row describing the dataset")
fname = input('Enter the fileName for the multicomponent data without the .txt and in
quotations:  ')
sample= input('Enter the fileName for the sample setup file without the .txt and in quotations:
')
prefix=input('Enter the new folder name to put analyzed data: ')
minT=input('Enter the minimum temperature of the file: ')
maxT=input('enter the maximum temperature of the file: ')
steps=input('enter the number of steps for the file: ')
offset=steps # can offset data manually if you like, or make it another input
print("Min-Max,Steps",str(minT)+"-"+str(maxT),steps)
cycleCol=input('which column, starting from 0, in the multicomponent file is the cycle
information in?: ')#column that cycle is in
print("you will be asked to enter fitting windows- this refers to the window for fitting baseline
correction ")
print("corrections will occur from ")
print fname,sample,prefix
rawSybr=input('If you would like rawfiles to output sybrFluorescence input 1, sybr/rox input
0: ')

if hideMe==1:
        photobleachCorrection=input('Would you like photobleaching correction? 1 for yes, 0
for no: ')
        if photobleachCorrection==1:
                default=input('would you like to use default tao and h parameters for
photobleaching? 1 for yes, 0 for no: ')
                if default==1:
                        tao= 7559*70/999 #The tao value for the exponential decay function
                        h = 0.93 # H value for exponential decay function
                else:
                        tao = input('please enter your tao value')
                        h = input ('please enter your h value')
```

```
	zeroBaseline=0  #set to 1 - will make all fluorescence below value of 0.0 have a value of 0.0
	print("you will be asked to enter fitting windows- this refers to the window for fitting baseline correction")
	print("for [[40,80],[45,85]] two sets of corrections will be performed, one with baselines from minT-40/80-maxT and the other from minT-45/85-maxT")
	fittingWindows=input('please input your fitting windows in sets of brackets- [[?,?]] for single sets, [[?,?],[?,?]] for multiple windows: ')
	test=1; asymptote=0.0;quantOffset=0.2#
	print("please enter the anneal/melt directions.  1 represents anneal, 2 represents melt. [1,2,1] would be a data set consisting of an anneal from MaxT to minT, followed by a melt and another anneal: ")
	direction=input('please enter your anneal/melts: ') # direction will set the melt/anneal direction.  All data will be re-arranged to go from 25-95,  1 is anneal, 2 is melt
	numStages=len(direction)
	print(direction,numStages)
	stageSet=[]#offset, offset+steps,offset+steps*2,offset+steps*3]
	blankStages=[]
	first=1
	blank=[]
	asymptoteCount=[]
	asymLoc=[]
	whichSet=input('would you like your final data from a linear baseline correction (1), exponential (2), or raw(3)?')#1 is linear, 2 is exponential, 3 is raw
	initialGuess=sy.array([1,-0.0104,592])#,0.945])  # convert parameter guess to something reasonable for data fitting
	parameterBounds=sy.array([(None,None),(-2,2),])
	for n in range(numStages): # for the number of stages, make an array for blank systems
		stageSet.append(offset+n*steps)
		blankStages.append([0,0,0,0,0])
		blank.append([])
		asymptoteCount.append(0)
		asymLoc.append([])
	print(stageSet)
	stages=[0,1,2]
	numStages=3
# Get sample data from setup file and make blank lists to hold appropriate data
if hideMe == 1:				# Data for nameset is organized by ["stringSampleName",[well1,well2,well3],[[well1],[well2],[well3]]]  where wellN =[stage1,stage2,stage3] where stage=(m,b,r^2,roxMax,sybrMax)]]]
	nameSet=[]; setup=open(sample+".txt",'r'); line=setup.readline() ; first=1 ; count=0
	for line in setup:
		if len(line.split()) >3:
			if first:
				split=line.split()
				print(line)
```

```
                                name=split[2] ; well=int(split[0])
                                nameSet.append([name,[well],[copy.deepcopy(blankStages)]])
                                first=0
                        else:
                                split=line.split()
                                name=split[2] ; well=int(split[0]) ; new=1 # is this a new name
                                for thisSample in nameSet:
                                        if thisSample[0]==name:
                                                thisSample[1].append(well)

        thisSample[2].append(copy.deepcopy(blankStages))
                                                new=0
                                if new==1:

        nameSet.append([name,[well],[copy.deepcopy(blankStages)]])
                        count=count+1
        setup.close()
        for n in nameSet: # print all the sample names
                print(n[0],n[1])
        avgSet=[];      dataSet=[];     index=0         #initialize blank data set -
        for n in range((96)+1):
                dataSet.append(copy.deepcopy(blank))
                index=index+1
if verbose==1: print(stageSet)

if hideMe==1:                           #Fill the dataset list with the fluorescence values from
the raw data file- for melt fill in the array backwards so all data runs 25-95C
        fIn=open(fname+".txt",'r') ; line= fIn.readline() ; thisRamp=[] ; maxCycle=0 ; index=0
        for line in fIn:
                splitLine=line.split()
                thisWell=int(splitLine[0].replace(',',''))
                if thisWell==1:
                        maxCycle+=1
                thisCycle=int(splitLine[cycleCol]) ; thisStage=0 ;
index=copy.deepcopy(thisCycle)
                for n in stageSet: #figures out which stage its in and sets the index
appropriately
                        if thisCycle > n:
                                thisStage+=1
                                index=index-steps # gives the index, as cycle goes up to 3,000
-- each stage only goes to 999
                if thisStage>numStages-1: #this only happens if there's a weird error
                        print "stageTooBig - this is an unusual error",thisCycle,thisWell
                if len(splitLine)<=3:  #if there's no data in this line
                        continue
                if thisStage>2:
                        print thisStage
```

```
                print thisCycle
            thisDirection=direction[thisStage]
            if thisDirection ==1: #anneal
                thisTemp=maxT - index*(maxT-minT)/steps
            elif thisDirection==2: #melt
                thisTemp=minT+index*(maxT-minT)/steps
            thisRox=float(splitLine[cycleCol+1].replace(',',''))
            thisSybr=float(splitLine[cycleCol+2].replace(',',''))
            thisNormal=copy.deepcopy(float(thisSybr/thisRox))
            if zeroBaseline ==1 and thisNormal < 0.0:
                thisNormal = 0.0
            if thisDirection ==1:

        dataSet[thisWell][thisStage].append([thisTemp,thisNormal,thisRox,thisSybr])
            elif thisDirection==2:

        dataSet[thisWell][thisStage].insert(0,[thisTemp,thisNormal,thisRox,thisSybr])
        fIn.close()

if hideMe==1:                       #Make the folders to put the data in
        maxSybr=0.00
        maxRox=0.00
        newFolder=1 ; stuffInThisFolder=os.listdir("./")
        for n in stuffInThisFolder:
            if n == prefix:
                newfolder=0
        if newFolder==1:
            os.mkdir(prefix)
            os.mkdir("./"+prefix+"/rawSybr")
            os.mkdir("./"+prefix+"/rawRox")
            os.mkdir("./"+prefix+"/rawNormal")
            os.mkdir("./"+prefix+"/hysteresis")
            for slopeCorrectionTemperature in fittingWindows:

        window=str(slopeCorrectionTemperature[0])+"_"+str(slopeCorrectionTemperature[1])
                os.mkdir("./"+prefix+"/"+str(window))

for n in nameSet:                   #Write the raw Rox data to the appropriate folder
        name="./"+prefix+"/rawRox/"+n[0]+"_raw.csv" ; fOut=open(name,'w') ;   titleLine=""
#fOut.write(n[0]) ;
        for m in (stageSet):
            titleLine=titleLine+",Temperature,"
            count=0
            for o in n[1]:
                titleLine=titleLine+"Replicate"+str(count)+","
        fOut.write(titleLine)
        index=0
```

```
	for m in range(steps):  #iterate the # of total cycles
		writeLine=[]
		first =1
		for thing in stages:
			writeLine.append("")
		z=2
		for o in n[1]: #o represents the well
			for thisStage in stages: #thing i s just an iterator for the current stage
				if first ==1:
					if index>=len(dataSet[o][thisStage]):
						writeLine[thisStage]=writeLine[thisStage]+",,"
#add block spaces between stages
					else:

	writeLine[thisStage]=writeLine[thisStage]+str(dataSet[o][thisStage][index][0])+","+str(
dataSet[o][thisStage][index][z])+","
					first=0
				else: # does not write the temperature
					if index>=len(dataSet[o][thisStage]):
						writeLine[thisStage]=writeLine[thisStage]+","
					else:

	writeLine[thisStage]=writeLine[thisStage]+str(dataSet[o][thisStage][index][z])+","
		first = 1
		index=index+1
		writeString=""
		for thisStage in stages:
			writeString=writeString+writeLine[thisStage]+","
		writeString=writeString+"\n"
		fOut.write(writeString)
	fOut.close()

for n in nameSet:			#Write the raw Sybr data to the appropriate folder
	name="./"+prefix+"/rawSybr/"+n[0]+"_raw.csv" ; fOut=open(name,'w') ;   titleLine=""
#fOut.write(n[0]) ;
	for m in (stageSet):
		titleLine=titleLine+",Temperature,"
		count=0
		for o in n[1]:
			titleLine=titleLine+"Replicate"+str(count)+","
	fOut.write(titleLine)
	index=0
	for m in range(steps):  #iterate the # of total cycles
		writeLine=[]
		first =1
		for thing in stages:
			writeLine.append("")
```

```
                z=3
                for o in n[1]: #o represents the well
                        for thisStage in stages: #thing i s just an iterator for the current stage
                                if first ==1:
                                        if index>=len(dataSet[o][thisStage]):
                                                writeLine[thisStage]=writeLine[thisStage]+",,"
#add block spaces between stages
                                        else:

        writeLine[thisStage]=writeLine[thisStage]+str(dataSet[o][thisStage][index][0])+","+str(
dataSet[o][thisStage][index][z])+","
                                        first=0
                                else: # does not write the temperature
                                        if index>=len(dataSet[o][thisStage]):
                                                writeLine[thisStage]=writeLine[thisStage]+","
                                        else:

        writeLine[thisStage]=writeLine[thisStage]+str(dataSet[o][thisStage][index][z])+","
                first = 1
                index=index+1
                writeString=""
                for thisStage in stages:
                        writeString=writeString+writeLine[thisStage]+","
                writeString=writeString+"\n"
                fOut.write(writeString)
        fOut.close()

for n in nameSet:                       #Write the raw Normal data to the appropriate folder
        name="./"+prefix+"/rawNormal/"+n[0]+"_raw.csv" ; fOut=open(name,'w') ;
titleLine="" #fOut.write(n[0]) ;
        for m in (stageSet):
                titleLine=titleLine+",Temperature,"
                count=0
                for o in n[1]:
                        titleLine=titleLine+"Replicate"+str(count)+","
        fOut.write(titleLine)
        index=0
        for m in range(steps):  #iterate the # of total cycles
                writeLine=[]
                first =1
                for thing in stages:
                        writeLine.append("")
                z=1
                for o in n[1]: #o represents the well
                        for thisStage in stages: #thing i s just an iterator for the current stage
                                fluor=dataSet[o][thisStage][index][z]
                                temp=dataSet[o][thisStage][index][0]
```

```
                                        if photobleachCorrection==1:

                fluor=fluor/modelFunctions.photoBleach2(temp,tao,h,direction[thisStage])
                                        if first ==1:
                                                if index>=len(dataSet[o][thisStage]):
                                                        writeLine[thisStage]=writeLine[thisStage]+",,"
#add block spaces between stages
                                                else:

                writeLine[thisStage]=writeLine[thisStage]+str(temp)+","+str(fluor)+","
                                                first=0
                                        else: # does not write the temperature
                                                if index>=len(dataSet[o][thisStage]):
                                                        writeLine[thisStage]=writeLine[thisStage]+","
                                                else:

                writeLine[thisStage]=writeLine[thisStage]+str(fluor)+","
                        first = 1
                        index=index+1
                        writeString=""
                        for thisStage in stages:
                                writeString=writeString+writeLine[thisStage]+","
                        writeString=writeString+"\n"
                        fOut.write(writeString)
                fOut.close()

#correct the data for decay
expSet=copy.deepcopy(dataSet)
for thisSample in nameSet:    #Write the raw data to the appropriate folder
        for thisWell in thisSample[1]:
                for thisStage in stages:
                        thisStageData=expSet[thisWell][thisStage]
                        minFluor=5.0
                        for thisPoint in thisStageData:
                                if thisPoint[1]<minFluor:
                                        minFluor=thisPoint[1]
                        for thisPoint in thisStageData:
                                thisPoint[1]-=minFluor-quantOffset
                        thisStageData.sort(key=itemgetter(0))
if photobleachCorrection:       #correct expSet for the decay function
        for thisSample in nameSet:
                stageCount=0
                for thisWell in thisSample[1]:
                        maxRox=0.00
                        maxSybr=0.00
                        for thisStage in stages:
                                minP=0.0
```

```
                                thisStageData=expSet[thisWell][thisStage]
                                numSteps=0
                                for thisPoint in thisStageData:
                                        if direction[thisStage] == 1:
                                                numCycles=thisStage*steps+numSteps+1.0
                                        else:
                                                numCycles=(1+thisStage)*steps-numSteps+1.0
                                        thisTemp=thisPoint[0]

        decayFit=modelFunctions.photoBleach2(thisTemp,tao,h,direction[thisStage])
                                        if decayFit== 1:
                                                print("dangit")
                                                print(numCycles,tao,h)
                                                print(float(numCycles)/tao)
                                                print(1/h)
                                                print(math.pow((numCycles/tao),(1/h)))
                                                sys.exit()
                                        thisPoint[1]=(thisPoint[1]/decayFit)
                                        numSteps+=1

if hideMe==1:                           #Perform the wittwer/quantum baseline corrections
        quantSet=copy.deepcopy(expSet)
        for thisSample in nameSet:
                wellIndex=0
                print thisSample[0],thisSample[1]
                for thisWell in thisSample[1]: # for each sample
                        if verbose==1: print thisWell
                        for thisStage in stages:# iterates through stages
                                if verbose ==1: print("thisStage ",thisStage)
                                initCurve= copy.deepcopy(quantSet[thisWell][thisStage])
                                transformCurve=[]
                                refFluor=0.0#
                                lowPoints=[]
                                highPoints=[]  # the higher section is the higher fluorescence,
not higher temp
                                refFluor=0.0
                                for thisPoint in initCurve:
                                        if thisPoint[1]>refFluor:
                                                refFluor=thisPoint[1]
                                for thisPoint in initCurve:
                                        initTemp=thisPoint[0]+273
                                        initFluor=thisPoint[1]
                                        transformTemp=-1.0/(25+273)+1.0/initTemp
                                        if (initFluor/refFluor) <= 0:
                                                print initFluor,refFluor,sy.log(initFluor/refFluor)
                                        transformFluor=sy.log(initFluor/refFluor)
                                        if thisPoint[0] <  slopeCorrectionTemperature[0]:
```

```
        lowPoints.append(copy.deepcopy([transformTemp,transformFluor]))
                                        if thisPoint[0] > slopeCorrectionTemperature[1]:

        highPoints.append([transformTemp,transformFluor])

        transformCurve.append([transformTemp,transformFluor])

                                xHighTemp=[];
        yHighTemp=[];xLowTemp=[];yLowTemp=[]#fluorescence
                                for p in highPoints: # sort into numpy arrays
                                        xHighTemp.append(p[0])
                                        yHighTemp.append(p[1])
                                for q in lowPoints:
                                        xLowTemp.append(q[0])
                                        yLowTemp.append(q[1])
                                xHigh= np.asarray(xHighTemp) ;
yHigh=np.asarray(yHighTemp)
                                xLow=np.asarray(xLowTemp) ; yLow=np.asarray(yLowTemp)
                                zHigh=np.polyfit(xHigh,yHigh,1) # line fit of x,y
                                zLow=np.polyfit(xLow,yLow,1) # line fit of x,y
                                funcHigh=np.poly1d(zHigh) ;
zHighTemporary=[];zLowTemp=[];    funcLow=np.poly1d(zLow)
                                if verbose==1: print "upper",zHigh[0],zHigh[1]," lower
",zLow[0],zLow[1]
                                tempCurve=[] #now normalize the data by the slope
                                for init,trans in zip(initCurve,transformCurve):
                                        transT= trans[0]
                                        fitHigh=funcHigh(transT)
                                        fitLow=funcLow(transT)
                                        fitHigh=sy.exp(fitHigh)*refFluor
                                        fitLow=sy.exp(fitLow)*refFluor
                                        thisFluor=sy.exp(trans[1])*refFluor
                                        if fitLow<fitHigh:
                                                correctedFluor=(thisFluor-fitLow)/(fitHigh-fitLow)
                                        else:
                                                correctedFluor=(thisFluor-fitHigh)/(fitLow-
fitHigh)
                                        if isinstance(correctedFluor,list):print "it's corrected"
                                        tempCurve.append([init[0],correctedFluor])
                                quantSet[thisWell][thisStage]=copy.deepcopy(tempCurve)
                        wellIndex=wellIndex+1

#perform line fitting *LINEAR* for photobleaching corrections
for slopeCorrectionTemperature in fittingWindows:
        print slopeCorrectionTemperature
        slopeSet=copy.deepcopy(expSet)
```

```
        for thisSample in nameSet: #for each sample
                if verbose==1: print(thisSample[0])
                wellIndex=0
                for thisWell in thisSample[1]:  #iterates through the wells
                        if verbose==1: print("thisWell ",thisWell)
                        for thisStage in stages:# iterates through stages
                                if verbose ==1: print("thisStage ",thisStage)
                                thisCurve= copy.deepcopy(slopeSet[thisWell][thisStage])
                                for dummy in stages:
                                        if verbose ==1: print("dataSet ",dummy," ",
len(dataSet[thisWell][dummy]),"expSet ",dummy," ", len(expSet[thisWell][dummy]))
                                if verbose==1: print("thisCurve ",len(thisCurve))
                                lowPoints=[]
                                highPoints=[]  # the higher section is the higher fluorescence,
not higher temp
                                maxP=0.0
                                minP=100
                                for p in thisCurve: # p is each data point - populates lists with
the fluorescence data
                                        if p[0] < slopeCorrectionTemperature[0]:
                                                highPoints.append(copy.deepcopy(p))
                                        if p[0] > slopeCorrectionTemperature[1]:
                                                lowPoints.append(copy.deepcopy(p))
                                        if p[0] > maxP: maxP=p[0]
                                        if p[0] < minP: minP=p[0]
                                xHighTemp=[];
        yHighTemp=[];xLowTemp=[];yLowTemp=[]#fluorescence
                                for p in highPoints: # sort into numpy arrays
                                        xHighTemp.append(p[0])
                                        yHighTemp.append(p[1])
                                for q in lowPoints:
                                        xLowTemp.append(q[0])
                                        yLowTemp.append(q[1])
                                if verbose==1: print("xHighTemp
",len(xHighTemp),"yHighTemp",len(yHighTemp),"xLowTemp ",len(xLowTemp),maxP,minP)
                                xHigh= np.asarray(xHighTemp) ;
yHigh=np.asarray(yHighTemp)
                                xLow=np.asarray(xLowTemp) ; yLow=np.asarray(yLowTemp)
                                zHigh=np.polyfit(xHigh,yHigh,1) # line fit of x,y
                                zLow=np.polyfit(xLow,yLow,1) # line fit of x,y
                                funcHigh=np.poly1d(zHigh) ;
zHighTemporary=[];zLowTemp=[];    funcLow=np.poly1d(zLow)

                                thisSample[2][wellIndex][thisStage][0]=[zHigh[0],zLow[0]]
                                thisSample[2][wellIndex][thisStage][1]=[zHigh[1],zLow[1]]
                                thisSample[2][wellIndex][thisStage][2]=[0,0]
```

```
                                thisCurve.sort(key=itemgetter(0)); tempCurve=[] #now
normalize the data by the slope
                                highOffset=0.0
                                lowOffset=0.0
                                asymptoteCorrection=0.0

                                for p in thisCurve:
                                        thisT=p[0]
                                        fitHigh=funcHigh(thisT)
                                        fitLow = funcLow(thisT)
                                        if p[1]>fitHigh and fitHigh>fitLow:
                                                if abs(p[1]-fitHigh)>abs(highOffset):
                                                        highOffset=p[1]-fitHigh
                                        fitLow=funcLow(thisT)
                                        if p[1]<fitLow:
                                                if abs(p[1]-fitLow)>abs(lowOffset):
                                                        lowOffset=p[1]-fitLow
                                highOffset=0.0
                                lowOffset=0.0
                                for p in thisCurve:
                                        thisT= p[0]
                                        fitHigh=funcHigh(thisT)
                                        fitLow=funcLow(thisT)
                                        thisFluor=p[1]
                                        correctedFluor=(thisFluor-fitLow)/(fitHigh-fitLow)
                                        tempCurve.append([thisT,correctedFluor])
                                slopeSet[thisWell][thisStage]=copy.deepcopy(tempCurve)
                        wellIndex=wellIndex+1

        #average the replicates - iterating in a slightly different order than before
        avgSet=[];avgSet2=[]
        thisSet=quantSet
        secondSet=slopeSet
        for thisSample in nameSet: # write out the hysteresis file
                averagedSample=[]
                averagedSample2=[]
                for thisStage in stages:
                        averagedStage=[]
                        averagedStage2=[]
                        for index in range(steps):
                                sumValue=0.0
                                sumValue2=0.0
                                for thisWell in thisSample[1]: #goes through replicate wells
                                        thisSet[thisWell][thisStage][index]

        sumValue=sumValue+thisSet[thisWell][thisStage][index][1]
```

```
        sumValue2=sumValue2+secondSet[thisWell][thisStage][index][1]
                                        temp=thisSet[thisWell][thisStage][index][0]
                                avgValue2=sumValue2/len(thisSample[1])
                                avgValue=sumValue/len(thisSample[1])
                                averagedStage.append([temp,avgValue])
                                averagedStage2.append([temp,avgValue2])
                        averagedSample.append([thisStage,averagedStage])
                        averagedSample2.append([thisStage,averagedStage2])
                avgSet.append([thisSample[0],averagedSample])
                avgSet2.append([thisSample[0],averagedSample2])
        window=str(slopeCorrectionTemperature[0])+"_"+str(slopeCorrectionTemperature[1])
        name = "./"+prefix+"/"+window+"/hysteresis.csv"
        fOut=open(name,'w')
        nameLine=""
        hystName=""
        for thisSample in avgSet:
                nameLine=nameLine+","+str(thisSample[0])+" anneal ,"+str(thisSample[0])+"
melt "
                hystName=hystName+","+str(thisSample[0])
        fOut.write(nameLine+","+hystName+"\n")

        for index in range(steps):
                writeLine=""+str(avgSet[0][1][thisStage][1][index][0])  #add the temperature
                hystLine=","
                for thisSample in avgSet:
                        if avgSet[0][1][thisStage][0] != thisStage:
                                print(avgSet[0][1][thisStage][0],thisStage," ug error")
                        writeLine=writeLine+","+str(thisSample[1][0][1][index][1])  #add the
normalized, averaged, theta stage 0
                        writeLine=writeLine+","+str(thisSample[1][1][1][index][1])  #add the
normalized, averaged, theta stage 1
                        writeLint=writeLine+","
                        hystLine=hystLine+","+str(abs(thisSample[1][0][1][index][1]-
thisSample[1][1][1][index][1]))
                fOut.write(writeLine+","+hystLine+", \n")
        fOut.close()
        for thisStage in stages: #write out the stage files
                name="./"+prefix+"/"+window+"/avgData_"+str(thisStage)+".csv"
                fOut=open(name,'w')
                nameLine=""

                for thisSample in avgSet:
                        nameLine=nameLine+","+str(thisSample[0])
                fOut.write(nameLine+"\n")

                for index in range(steps):
```

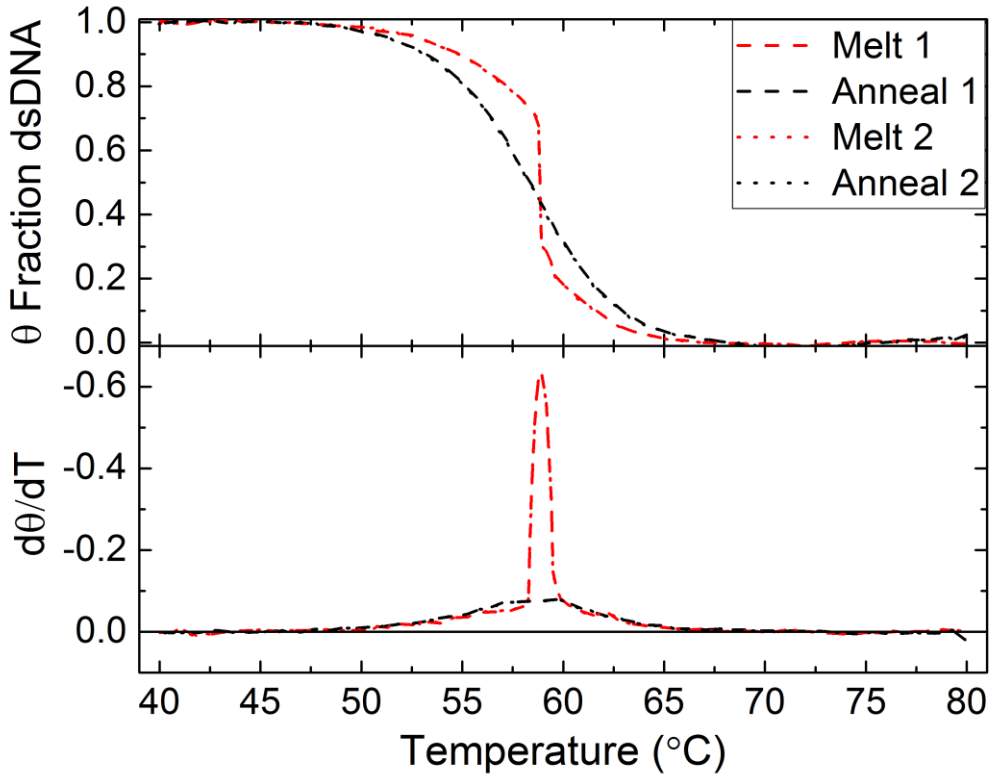
```
                        if avgSet[0][1][thisStage][0] != thisStage:
                                print(avgSet[0][1][thisStage][0],thisStage," ug error")
                        writeLine=""+str(avgSet[0][1][thisStage][1][index][0])  #add the
temperature
                        for thisSample in avgSet:

        writeLine=writeLine+","+str(thisSample[1][thisStage][1][index][1])  #add the
normalized, averaged, theta
                        if test==1:
                                for thisSample in avgSet2:

        writeLine=writeLine+","+str(thisSample[1][thisStage][1][index][1])  #add the
normalized, averaged, theta
                        fOut.write(writeLine+"\n")
                fOut.close()

sys.exit()
```

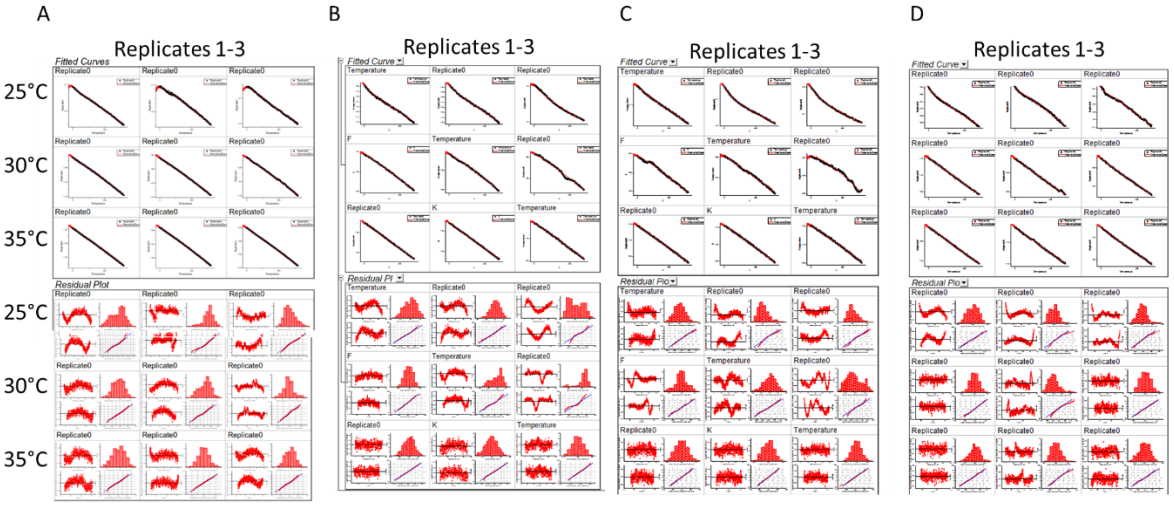
Appendix 2: Weave Tile UV-Vis Data



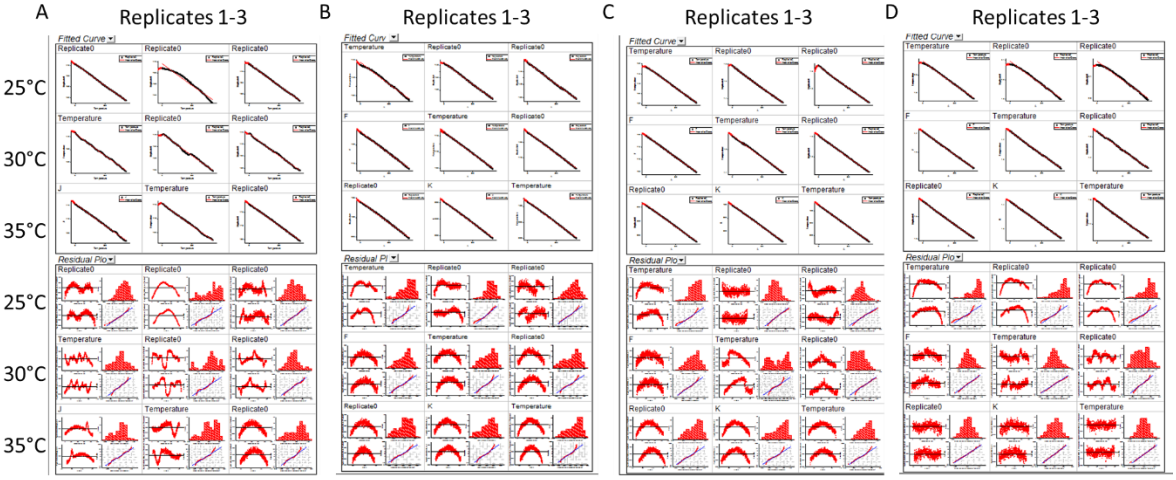
Weave tile UV-Vis data obtained by previous postdoc of the Weave Tile used in Chapter 3.

While the samples exhibit a small hysteresis, but with identical transition temperature.

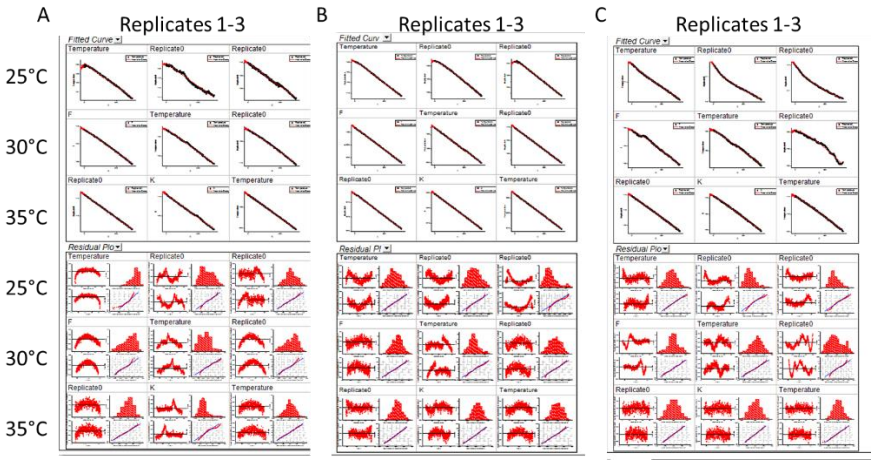
Appendix 3: Photobleaching Fitting



Fitting data for A- Tall Rectangle inTAE Mg buffer 5 nM pH 7.4, B- Tall Rectangle in Sodium Cacodylate Buffer Mg 5 nM pH 7.4, C- Tall Rectangle in Sodium Cacodylate Buffer Mg 5 nM pH 5.5, D- Tall Rectangle in Sodium Cacodylate Buffer Mg 2.5 nM pH 7.4
Columns are replicates of the same samples.
Rows are the three temperatures each occurred at.



Fitting data for A- Triangle in Sodium Cacodylate Mg buffer 5 nM pH 7.4, B- Bundle in Sodium Cacodylate Buffer Mg 5 nM pH 7.4, C- Circle in Sodium Cacodylate Buffer Mg 5 nM pH 5.5, D- Circle in Sodium Cacodylate Buffer Mg 2.5 nM pH 5.5
Columns are replicates of the same samples.
Rows are the three temperatures each occurred at.



Fitting data for A- Circle in TAE Mg buffer 5 nM pH 7.4, B- Circle in Sodium Cacodylate Buffer Mg 2.5 nM pH 7.4, C- Tall Rectangle in Sodium Cacodylate Buffer Mg 2.5 nM pH 5.5
Columns are replicates of the same samples.
Rows are the three temperatures each occurred at.

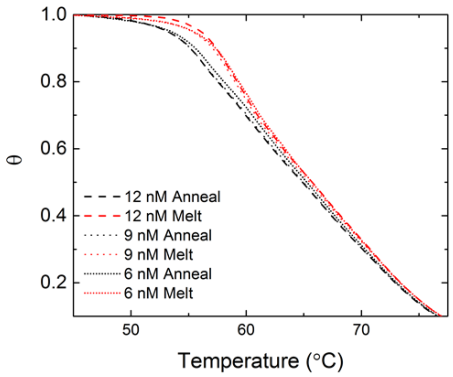
The top set of curves for each replicate/temperature pair is the fluorescence intensity and phobleaching fit data. The lower set of curves presents multiple representations of the residuals of the curve fit.
As seen in the overlap of the top curve set, and the relatively high contribution of random noise to the residual in the lower curve set indicate a good fit.

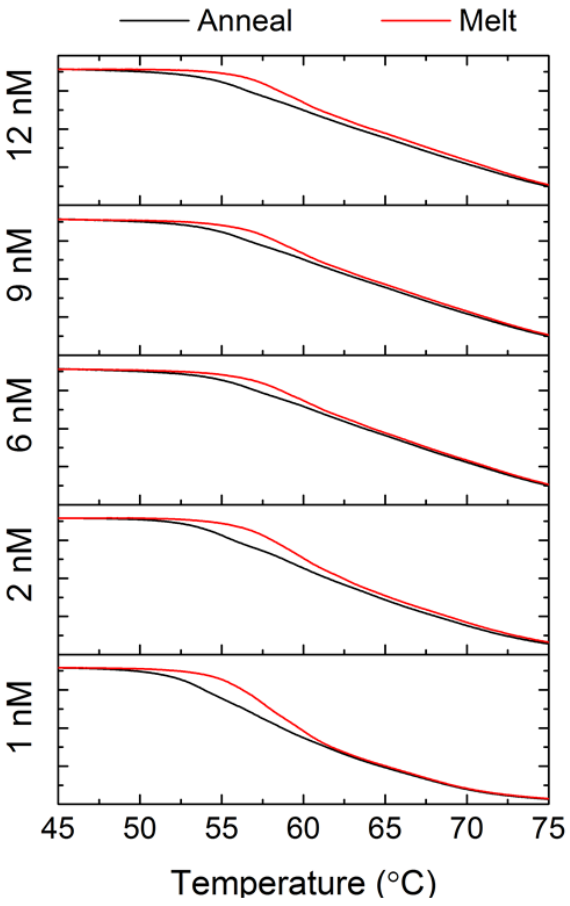
Appendix 4: Anneal/Melt data hysteresis

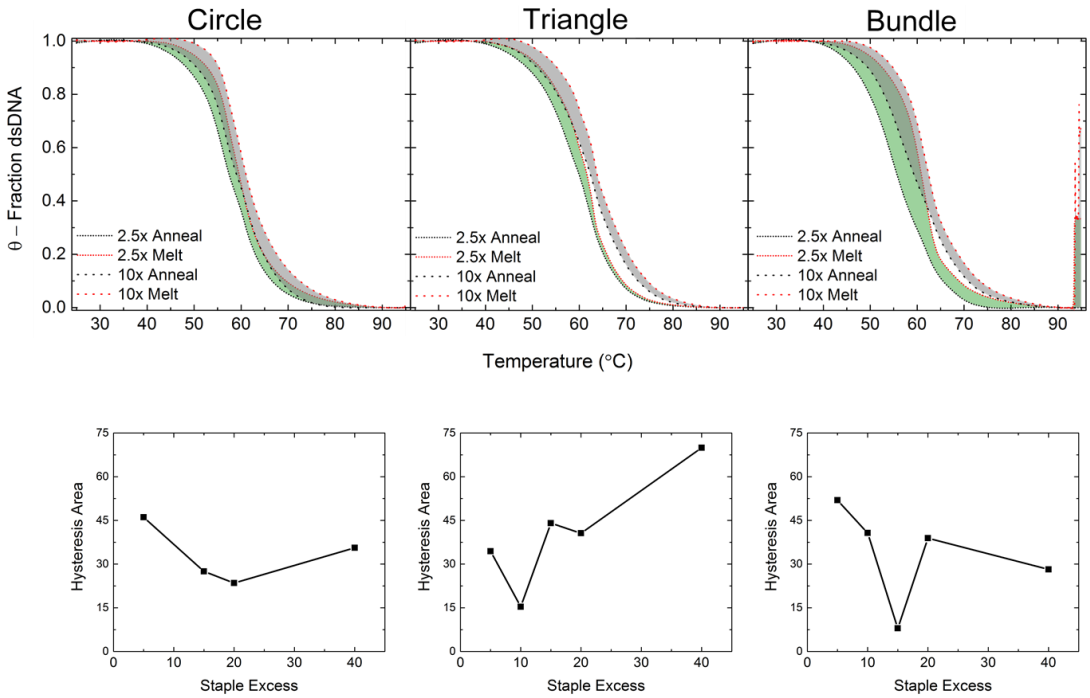
Left, overlay of hysteresis data for 6, 9, and 12 nM Tall Rectangle origami.

Melt curves appear to have a greater shift than anneal curves.

Below, anneal/melt curves separated by scaffold concentration.



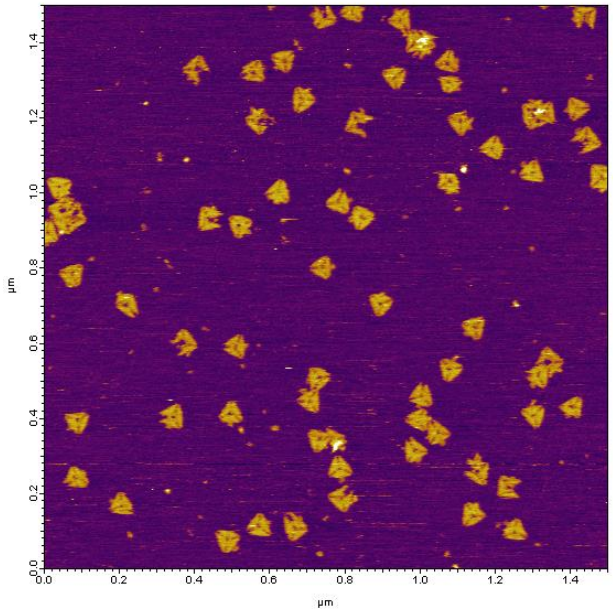




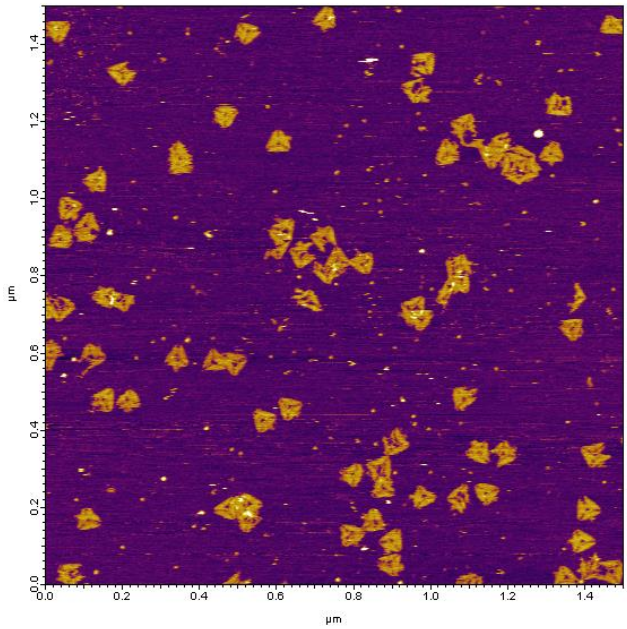
Hysteresis data as a function of staple excess for the miniM13 origami shapes.

Appendix 5: Triangle Binary Systems

The triangle structures showed clear dominance in competitive anneals. Below are examples of binary multishape anneals including the triangle shape. These had the lowest excess triangle, and looked identical to the systems which had a higher triangle excess, which is to say that all looked like pure triangle.



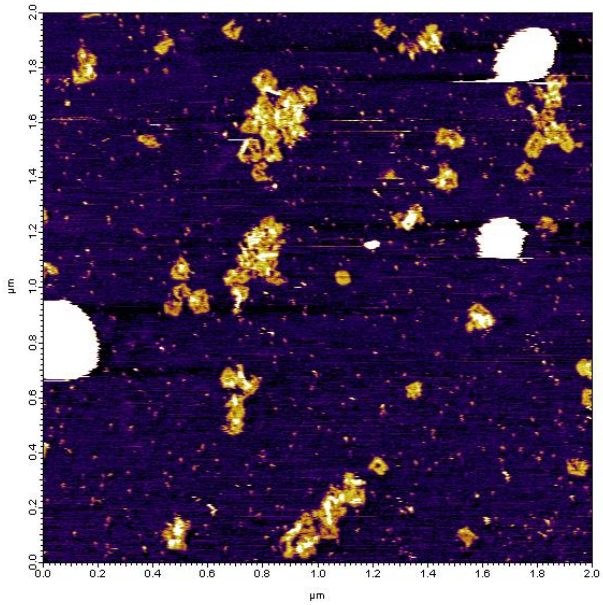
2.5xTriangle, 5xBundle



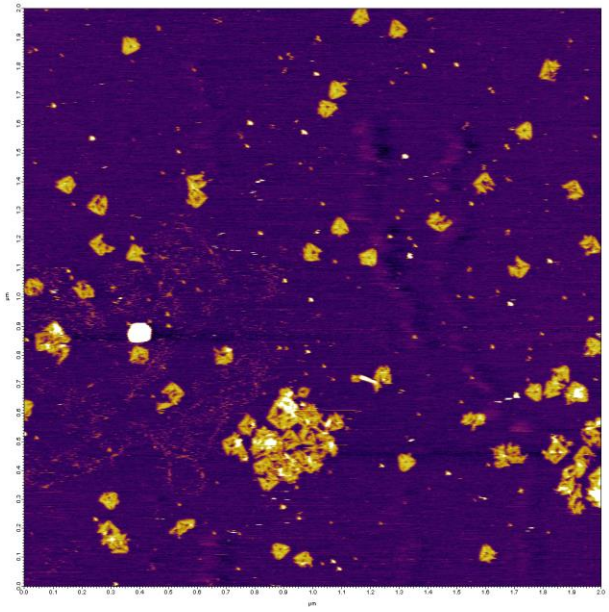
2.5x Triangle, 5xCircle
The circle had more aggregate type shapes, but still clearly showed triangle dominance

Appendix 6: Ternary Anneals

In Fig. 4.9, Ternary anneals were depicted. For clarity, these anneals omitted the large aggregates visible in the ternary Circle/Bundle/Triangle system at 7.5x staples. It is likely that a higher overall staple excess is needed to form chimera in a ternary system.



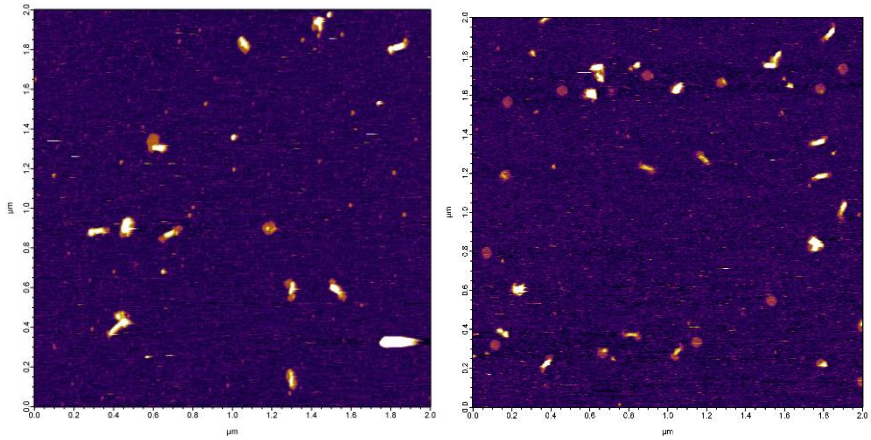
0.6xTriangle, 4.7xCircle, 4.7xBundle



2.8xTriangle, 3.6xCircle, 3.6xBundle

Appendix 7: ImageJ Analysis of Binary Systems

Example images of binary systems



25 nM 5xBundle/2.5xCircle 8 hr 50 nM 3.75xBundle/3.75xCircle 8 hr

Average percent bundle of each chimera					percent chimera – by count					corrected percent bundle			
2.5C-5B	**5 nM**	**25 nM**	**50 nM**		**2.5C-5B**	**5 nM**	**25 nM**	**50 nM**		**2.5C-5B**	**5**	**25**	**50**
2Hr	0.536881	0.531366	0.497874		**2Hr**	91.67 %	75.69 %	76.99 %		**2Hr**	49.21 %	40.22 %	38.33 %
8Hr	0.541569	0.539711	0.537853		**8Hr**	93.54 %	87.93 %	82.62 %		**8Hr**	50.66 %	47.46 %	44.44 %
16Hr	0.60929	0.652198	0.592129		**16Hr**	96.27 %	96.38 %	87.01 %		**16Hr**	58.66 %	62.86 %	51.52 %
3.75C-3.75B	**5 nM**	**25 nM**	**50 nM**		**3.75C - 3.75B**	**5 nM**	**25 nM**	**50 nM**		**3.75C - 3.75B**	**5 nM**	**25 nM**	**50 nM**
2Hr	0.296996	0.436866	0.399209		**2Hr**	24.00 %	36.59 %	36.59 %		**2Hr**	7.13%	15.98 %	14.61 %
8Hr	0.421203	0.43605	0.450896		**8Hr**	66.34 %	55.29 %	61.90 %		**8Hr**	27.94 %	24.11 %	27.91 %
16Hr	0.507502	0.473223	0.507502		**16Hr**	76.92 %	52.71 %	56.11 %		**16Hr**	39.04 %	24.95 %	28.48 %
5C-2.5B	**5 nM**	**25 nM**	**50 nM**		**5C-2.5B**	**5 nM**	**25 nM**	**50 nM**		**5C-2.5B**	**5 nM**	**25 nM**	**50 nM**
2Hr	0.32	0.322614			**2Hr**	4.55%	13.10 %			**2Hr**	1.45%	4.23%	
8Hr	0.353574	0.345101	0.348467		**8Hr**	31.25 %	46.15 %	23.02 %		**8Hr**	11.05 %	15.93 %	8.02%
16Hr	0.382411	0.402076	0.374965		**16Hr**	31.25 %	22.53 %	9.79%		**16Hr**	11.95 %	9.06%	3.67%

Table of Data for counts and corrected percentages plotted in chapter 4